Essays on Pension Scheme Design and Risk Management

Jiajia Cui

Essays on Pension Scheme Design and Risk Management

PROEFSCHRIFT

terverkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op vrijdag 16 januari 2009 om 10.15 uur door

JIAJIA CUI

geboren op 30 juni 1977 te Zhangjiakou, China

PROMOTORES
prof. dr. Frank de Jong
prof. dr. Eduard Ponds

PROMOTIECOMMISSIE
prof. dr. Alexander Michaelides
prof. dr. Bertrand Melenberg
prof. dr. Coen Teulings
prof. dr. Bas Werker

Contents

Acknowledgements **v**

1 Introduction **1**

2 Intergenerational risk sharing within funded pension schemes **7**

2.1 Introduction . . . 7
2.2 Economy, overlapping generations and preferences . . . 11
2.3 Pension schemes with intergenerational risk sharing . . . 13
2.3.1 Pension liability . . . 13
2.3.2 Risk allocation rules . . . 15
2.3.3 Pension policies optimization . . . 17
2.4 Evaluation of pension schemes . . . 17
2.4.1 Optimal individual scheme benchmark . . . 18
2.4.2 Optimal collective schemes and welfare evaluation . . . 20
2.4.3 Welfare of older and future generations . . . 23
2.4.4 Sensitivity analysis . . . 24
2.5 Sustainability . . . 25
2.5.1 Market valuation of the pension schemes . . . 26
2.5.2 Initially underfunded and overfunded schemes . . . 28
2.5.3 Social planner's view . . . 28
2.6 Conclusion . . . 30
2.7 Appendix A: Solution method of collective schemes . . . 30
2.8 Appendix B: Solution method of individual schemes . . . 31
2.8.1 OI scheme . . . 31

2.8.2 DC scheme 32
2.8.3 Sensitivity analysis 33
2.9 Tables and Figures 35

3 DC pension plan defaults and individual welfare 49
3.1 Introduction 49
3.2 Model Setup 54
3.2.1 Individual's preferences 54
3.2.2 Return dynamics 54
3.2.3 Labor income and social security 55
3.2.4 Housing and medical expenditures 56
3.2.5 Tax-Deferred Account and Taxable Account 57
3.3 Optimal life cycle strategies 58
3.3.1 Optimization problem with TDA and TA 58
3.3.2 The life cycle saving and investing profiles 60
3.3.3 With Employer Matching 61
3.3.4 Welfare evaluation 62
3.4 Default designs 62
3.4.1 Default #1: constant contribution rate and constant portfolio 63
3.4.2 Default #2: constant contribution rate and age-dependent portfolio 64
3.4.3 Default #3: age-dependent contribution rate and constant portfolio 65
3.4.4 Default #4: age-dependent contribution rate and age-dependent portfolio 66
3.5 Welfare comparisons 67
3.5.1 Baseline model 67
3.5.2 Sensitivity analyses 69
3.6 Conclusions 72
3.7 Appendix A: Solution method of Section 3, with TA and TDA . . 74
3.7.1 Solving the retirement period 74
3.7.2 At time $t = R - 1$ 76
3.8 Appendix B: Solution method of Section 4 83

3.9 Tables and Figures 87

4 Longevity Risk Pricing 101

4.1 Introduction 101
4.2 Stochastic Mortality Models and Longevity-Linked Securities . . . 105
4.2.1 Stochastic Mortality Models 105
4.2.2 Longevity-Linked Securities 109
4.3 Longevity Risk Pricing Principles, A Review 110
4.3.1 An Application 113
4.3.2 Preference Assumption 117
4.4 Pricing of a Coupon-Based Longevity Bond 117
4.4.1 Setup 118
4.4.2 Results 119
4.4.3 Implications 122
4.5 Pricing of Other Longevity-Linked Securities 122
4.5.1 Vanilla Longevity Swaps 122
4.5.2 Deferred Longevity Bonds 123
4.5.3 Longevity Floors and Longevity Caps 125
4.6 The Effect of Natural Hedging 126
4.7 The Demand Side Pricing and Basis Risk 129
4.7.1 Demand Side Pricing 129
4.7.2 Basis Risk 130
4.8 Conclusion 133
4.9 Appendix A: The Lee-Carter 1992 Model 134
4.9.1 Estimation Procedure of LC92 Model 135
4.9.2 Simulation 136
4.10 Appendix B: Derivation of the Results in Section 4 137

References 141

Nederlandse Samenvatting (Dutch Summary) 149

Acknowledgements

This PhD thesis is the fruit of three years of research at the Department of Finance at Tilburg University, and two years of graduate study at Tinbergen Institute Amsterdam. I would like to take this opportunity to thank a number of people and organizations that have made this thesis possible and enjoyable. I also acknowledge All Pensions Group (APG) for its generous financial support for my PhD research.

I am particularly grateful to my thesis supervisors, Frank de Jong and Eduard Ponds, for their invaluable guidance and support throughout the past five years. I have been extremely fortunate to work with them and benefit from their intensive expertise and merits. Frank and Eduard have played an important role in my life and my career. They have spent a great deal of time with me, trained me to be an academic researcher, and advised me on my career choices. They have also provided me with an excellent research environment in which I could combine my academic research with the leading pension practices at APG.

Frank has taught me many of his merits and skills, which are invaluable for many aspects of life. "Learning-by-doing" is how Frank encouraged me six years ago, when I started my Master's thesis with him. Since then, I have never been afraid to explore a new area. During my PhD research, Frank gave me a great deal of freedom to define and develop my research interests. (Such freedom should not be taken for granted, as many PhD students have to follow pre-defined topics.) Defining my own research topic was a painful but most beneficial process. Frank was very critical about my proposals. His critical but constructive comments helped me to sharpen my arguments and identify a potential contribution to the literature. "Be very focused" and "do one thing at a time", are the keys to success, as Frank keeps reminding me. Among his other merits, his careful preparation and excellent time-management skills are still my learning points for the future. Thank you Frank!

Eduard has been another great supervisor, as he always promotes young people and has inspired me with his endless new ideas. I admire his great enthusiasm for his work and his creativity in the area of pension reform. Eduard brought me the big challenge of "designing the optimal pension scheme", which is the central theme of my thesis. Although we have worked on several specific aspects of this issue during my PhD research, I still have a long way to go before I reach this ultimate goal. I feel honoured to be able to continue working with Eduard on this innovative project in years to come. During the course of my PhD research, Eduard introduced me to many pension experts at APG, so that I have been exposed to leading pension fund practices and, more importantly, to the frontier of the development of pension industry. These contacts have helped me to define my research questions with the aim of solving practical problems. It was definitely helpful for writing papers and obtaining a research grant. For all of your inspiration and encouragement, thank you Eduard!

My special thanks go to my committee members, Alexander Michaelides, Bas Werker, Bertrand Melenberg and Coen Teulings. Alex hosted me for two months at LSE (Financial Markets Group) when I was working on my job market paper. He made time available for me, commenting on my research and suggesting directions for future research. Bas is well known for his quick understanding and critical remarks. I benefited greatly from the critical comments and discussions of Bas and Frank during the finance seminars, which pinpoint the kernels of the issues. I am grateful to Bertrand and Coen for their detailed and constructive comments on my dissertation.

I feel privileged to visit the Financial Market Group at LSE in 2007 for two months, and have met so many nice people and to have been able to discuss my research at this top research institute. I have benefited a great deal from many interesting and insightful discussions with David Webb, Alex Michaelides, Christian Julliard, Ron Andersen and Paula Lopes. I also thank my officemates Yuki Sato and Sheng Li, who made my stay at LSE very pleasant. The generous financial support of CentER the funding is also appreciated.

I would also like to thank the Finance Department, Netspar, CentER and

Tinbergen Institute for creating a stimulating research environment. PhD research is often characterized as a lonely journey. But, as a member of Netspar, I feel that my research has been part of a big team effort. Netspar is a broad research network focusing on pension- and related studies. Through this network, I am connected to Dutch as well as international scholars and practitioners for sharing knowledge and discussing those grand issues that are of concern to ageing societies. I have benefited enormously from participating and presenting at Netspar seminars, workshops and conferences. I am therefore grateful to the founders of Netspar, Lans Bovenberg and Theo Nijman, and a number of Netspar researchers who kindly spent time with me and commented on my research. I am especially grateful to Norma Coe, Katie Carman, Joachim Inkmann, Paul Sengmuller, Hans Schumacher, Johannes Binswanger, Joost Driessen (UvA) and Antoon Pelsser (UvA), Berend Roorda (UTwente), Michel Vellekoop (UTwente), Nico Mol (UTwente) and Arun Bagchi (UTwente).

I am also indebted to a number of colleagues at APG for taking an interest in my research and constantly inspiring me with new challenges in practice. Particularly, I would like to thank Roderick Molenaars, Rob van der Goorbergh, Onno Steenbeek, Peter Vlaar, Roy Hoevenaars, Karen Aarssen, Henk Angeman, Johannes Krottje and Stefan Lundbergh. I am glad that I will be able to continue working with you after completion of my PhD research on many challenging pension issues.

I feel privileged to have so many supportive friends who have made my life at the University of Amsterdam, Tinbergen Institute and Tilburg University so very enjoyable. I have greatly benefited from my previous colleagues and good friends at UvA: Otto van Hemert, David Schrager, Roger Laeven and Yiu Chung Cheung. For many years they have provided me with good advice for carrying out research and making career choices. From Tinbergen Institute, I particularly would like to thank Valeriu Trufas, Sumedha Gupta, Tse-chun Lin, Dion Bersem, Marcel Vorage, Ronald Wolthoff, Ebby Bekkers, Jiang Huang, Yin-yen Tseng, Cocco Huang, Lei Pan, Michel van der Wel and Razvan Vlahu, for their great help from time to time. From Tilburg University, I appreciate the kind help I have received from Zhen Shi, Ralph Koijen, Roel Mehlkopf, Viorel Roscovan, Kim Peijenburg, Ilky Sendeniz-Yuncu, Thijs van der Heijden, Jingqiang Guo, Esther

Eiling, Jeremie Lefebvre, Corina Pungulescu, Marta Szymanowska, Chendi Zhang, Renxiang Dai, Jie Zheng, Karen van der Wiel, Peter Szilagyi, Igor Loncarski, Norbert Hari, Marie-Cecle Kwinten, Nicole Segers, Loes de Groot, Sharon Lueb. Also, special thanks to my kind housemates Ting Jiang and Yang Zhao. I am grateful to Jeanne Bovenberg-Meyer for her meticulous editing of a part of my thesis, and to Roel for editing the Dutch summary.

Finally, my deepest thanks go to my parents and parents-in-law, my lovely sister Jingjing and my dear husband Jianbin, for their unconditional support, courage and faith in me.

Jiajia Cui
Tilburg, September 2008

Chapter 1

Introduction

Funded pension systems have been evolving as the major source for retirement provision internationally. The ageing process and the new accounting regulation of market-based valuation of liabilities, however, have caused the costs of financing traditional DB (defined benefit) schemes to escalate dramatically— and it has become a struggle for public and private sponsors to fulfill these funding commitments. Particularly during the market meltdown of 2001-2003, many DB pension funds either closed their doors to new members or froze up completely. Funded pension schemes worldwide have been pushed toward a crossroad of reform.

In recent years, two directions for funded pension systems have been taken. The first entails abolishment of the collective arrangement, and a direct move into individual pension schemes, where funding responsibility is shifted completely from sponsors to individuals. Various individual defined-contribution (DC) accounts are typical examples here. However, the DC schemes not only concentrate the risks on each individual, but also confront individuals with complex investment decisions. The second direction involves keeping the collective nature of the DB schemes, but spreading out the funding responsibility to all stakeholders (retirees, employees and sponsors). Various hybrid collective schemes are typical examples of such practices. It puts forward the question as to how risks should be allocated among stakeholders. Is one direction better than the other, or should both directions be improved? These are the questions that this thesis attempts to answer.

This thesis contributes to the economic understanding of the relative strengths and weaknesses of both collective and individual pension schemes, and puts forward several proposals for further improvement of funded pension systems. In a nutshell, the economic added value of collective schemes lies in efficient risk sharing, which is the weakness of individual schemes; The strength of individual schemes is the potential tailor-making to heterogeneous preferences. The next generation of pension schemes should therefore combine the strengths of both worlds, namely the intergenerational risk-sharing property of the collective schemes and the life-cycle characteristics of the individual schemes. The detailed arguments are presented in the following chapters.

Chapter 2 ("*Intergenerational risk sharing within funded pension schemes*", with Frank de Jong and Eduard Ponds) of this thesis focuses on one unique feature of the collective schemes, which is the possibility of intergenerational risk sharing (IRS) among workers and retirees covered in the same pension scheme. The key intuition is that, the long term nature of the collective pension funds allows for smoothing of risks across many generations. The collective schemes with IRS thus provide broader risk sharing possibilities and mutual insurance benefit, as compared with individual schemes, which concentrate the (financial and non-financial) uncertainties on single individual. We demonstrate that the well organized intergenerational risk sharing mechanism within realistic collective funded schemes can be welfare improving above the theoretical optimal individual scheme benchmark. This new finding has profound implications for the pension reforms for both collective and individual systems. The intergenerational risk sharing property of collective schemes should be retained and strengthened, in stead of being abolished, during the reforms. Our case study shows that IRS is implementable in realistic pension funds.

Chapter 2 is closely related to three literature areas, namely the Asset Liability Management (ALM), Intergenerational Risk Sharing, and contingent claim analysis. The traditional ALM framework takes the DB pension (contribution and benefit) policies as given, and focuses on the strategic portfolio allocation decisions.[1] Our approach is different in that, we integrate the design of contribution

[1] The insightful literatures along this line include Sharpe and Tint (1990), Campbell and Viceira (2002), Hoevenaars, Molenaar, Schotman and Steenkamp (2007), Van Binsbergen and Brandt (2007).

and benefit policies into the ALM framework; therefore, the pension policies (i.e. risk sharing rules) are optimized together with asset allocation decisions. We show that pension policy designs affect asset allocation choices as well as the welfare of the participants. Building on the literatures on intergenerational risk sharing (IRS) which mainly focus on the public finance,[2] we analyze whether IRS is desirable in funded pension schemes and if so, what the optimal design regarding risk sharing rules is. Furthermore, our valuation framework in Chapter 2 is consistent with the market-based valuation of pension and insurance liabilities.[3] With these contingent claim valuation techniques, we are able to decompose the market value of the intergenerational risk sharing into call and put options held by each generation. We show that, ex ante, the market value of the call option equals the market value of put option; however, the IRS may lead to welfare improvements for the participants of collective schemes.

However, the collective schemes with IRS are not perfect either. One particular criticism is that its uniform policies do not customize to the heterogeneous profiles of all participants. However, tailor making to each individual's profile is exactly the strength of (ideal) individual schemes. Chapter 3 therefore examines the individual DC scheme design in a realistically calibrated life cycle framework.

The study of individual DC schemes in **Chapter 3** ("*DC Pension Plan Defaults and Individual Welfare*") serves two purposes: First, the resulting optimal life-cycle contribution and investment strategies provide useful information for the re-design of collective schemes, for example, incorporating the life cycle profile into collective pension policies. This leads to a new work-in-progress on (w)age-dependent policy design for collective pension schemes. Second, the resulting optimal life-cycle strategies provide useful indications for designing the default settings of the individual schemes. Theoretically, the default design does not matter, as each participant has the freedom to choose and implement his/her optimal

[2] Gordon and Varian (1988) and Shiller (1999) give a general exposition of these issues. Merton (1983), Enders and Lapan (1982), Krueger and Kubler (2006), and Van Hemert (2005) discuss the IRS in social security programs; Allen and Gale (1997), Van Bommel (2006), and Gollier (2007) discuss the IRS in financial intermediaries.

[3] Related literatures on the valuation of pension liabilities, guarantees and embedded options in insurance contracts include Grosen and Jorgensen (2000), Schrager and Pelsser (2004), De Jong (2008a).

strategies. However, in reality, we observe that most people simply follow the given defaults, which are often constant. Given the dramatic impact of default designs, Chapter 3 therefore proposes age-dependent contribution and investment default designs. We show that the simple age-dependent defaults dramatically improve the welfare of the participants (above the current constant default policies).

Chapter 3 builds on the literature on life cycle consumption and portfolio choice.[4] This chapter closely relates to a paper by Gomes, Michaelides and Polkovnichenko (2008), who study the optimal portfolio choices and optimal saving strategies for rational individuals with a taxable account and a tax-deferred DC account. Chapter 3 adapts a similar modeling setup, but with a different focus on the optimal age-dependent contribution and investment default rules. Gomes, Kotlikoff and Viceira (2008) also study the welfare comparisons of simple defaults. However they only consider defaults for portfolio choice. We find that the contribution (or saving) default has larger impact on welfare than the portfolio choice does.

One simplification we made in Chapter 3 is the annuitization strategies of the individuals.[5] Although we simplified the annuitization issues at the individual level, we devote Chapter 4 on the longevity risk management issues at the aggregate level. As we know that, longevity trend and its uncertainties drive the global ageing and have triggered the massive reforms in the pension world. Both collective pension funds and annuity providers associated with individual DC schemes are subject to longevity risk. In several notable cases, it leads to severe solvency issues for pension funds and annuity providers.

One effective way of managing longevity risk is by transferring longevity risk to the broad financial market via the so-called longevity linked securities. But longevity-linked securities are not traded in financial markets due to the pricing difficulty. **Chapter 4** ("*Longevity Risk Pricing*") proposes a new method to price

[4] Insightful literature on this line include Merton (1969, 1971), who emphasizes on the role of human capital, Carroll (1992, 1994, 1997) on precautionary saving, Gourinchas and Parker (2002) on life cycle consumption, Viceira (2001), Gomes and Michaelides (2005) and Cocco, Gomes and Maenhout (2005) on life cycle portfolio choice.

[5] We refer to recent papers by Davidoff, Brown, and Diamond (2005), Koijen, Nijman, and Werker (2007), Horneff, Maurer, Mitchell, and Stamos (2008) for various aspect of annuity markets and annuitization issues.

the longevity risk premia in order to tackle the pricing obstacle under incomplete market setting. We show that the size of the risk premium depends on the payoff structure of the security due to the market incompleteness. Furthermore, we show that the financial strength of the longevity insurance seller and buyer, the availability of the natural hedges and the presence of basis risk may significantly affect the size of longevity risk premium.

Chapter 4 combines the literatures on incomplete market pricing principles[6] with the literatures on stochastic mortality modeling[7], in order to tackle the pricing difficulty of the innovative longevity linked securities.

Based on these findings, this thesis also suggests several directions for future improvements of pension schemes. The next generation of pension schemes might combine the strengths of both collective and individual schemes. How can we do that? For instance, the DC schemes could add profit and risk sharing components to replicate IRS. The collective schemes could incorporate life cycle profiles in their policies. To sum up, with the help of pension engineering and financial market innovations discussed in this thesis, the sound retirement security is possible in ageing societies.

[6] The related literatures on incomplete market pricing include papers by Cochrane and Saa-Requejo (2000), Svensson and Werner (1993), Young and Zariphopoulou (2002), Musela and Zariphopoulou (2004), De Jong (2008).

[7] The related literatures on stochastic mortality modeling include papers by Lee and Carter (1992), Hari, De Waegenaere, Melenberg, and Nijman (2008a, b), Schrager (2006), Cairns, Blake, and Dowd (2005).

Chapter 2

Intergenerational risk sharing within funded pension schemes

This chapter is based on Cui, De Jong and Ponds (2008).

2.1 Introduction

Systematic risks cause dramatic and long-lasting up- and down-swings in the economy and impose significant impacts on millions of households.[1] As a market equilibrium outcome, in face of financial crisis, consumptions have to be sharply reduced for a significant period of time.[2] Allen and Gale (1997) argue that there are two strategies which can improve intertemporal smoothing of risks, one is intergenerational risk sharing via public programs and the other one involves asset accumulation via financial intermediaries. Indeed, the private market fails to provide insurance products based on intergenerational risk sharing because human capital in non-tradable and current generations cannot sign contracts with future generations. We examine an important question missing in this literature, namely whether intergenerational risk sharing (IRS) is desirable in funded pen-

[1] The credit crunch since 2007, the internet bubble in the beginning of this century, the lost decade in Japanese market, the oil crisis in 1970s, and the great depression in 1930s are typical examples of such systematic risks.

[2] See also Teulings and De Vries (2006) for an illustration.

sion schemes with mandatory participation, and if so, what the optimal design regarding risk sharing rule is. Our hypothesis is that, collective funded pension schemes with IRS, due to its long-term nature, could be used as dedicated financial institutions to facilitate intertemporal smoothing of systematic risks. This paper extends the literature on the role of intergenerational risk sharing within funded pension schemes on intertemporal smoothing. Our study provides policy implications regarding the on-going reforms of pension systems worldwide.

In this paper, we focus on intergenerational sharing of systematic investment risk in realistic funded pension schemes. The schemes analyzed are stylized versions of collective (public sector, industry-wide, or multi-employer) pension funds operating in countries like the US, UK, Canada and the Netherlands. In these collective schemes, assets are pooled and owned collectively by young and old generations. Mandatory participation is required by legislation. The schemes invest in the financial markets, typically partly in risky assets (stocks) to capture the equity premium. The schemes have no sponsor and investment risk is borne by the overlapping generations of pension plan members. Surpluses or deficits in the funding process are shared among young, old and future generations by adjusting either contributions, benefit levels or both, which leads to intergenerational transfers. Under this setup, we carry out a welfare comparison of various collective schemes with IRS, and use the optimal individual scheme as the benchmark.

More specifically, this paper specifies a number of realistic collective pension schemes consisting of 55 overlapping generations (from age 25 to 80). During the first 40 years workers contribute to the pension fund and the last 15 years retirees receive benefits. Participants are short-sales and borrowing constrained, and their only savings are for retirement purposes. The pension funds invest the savings of all cohorts in one asset pool, and part of it may be invested in the stock market. The return on the risky asset is exogenous and unaffected by the scheme's design or portfolio composition. Fluctuations in asset returns lead to mismatch risk between the pension fund assets and liabilities. Contributions and/or benefits are adjusted as a function of fund surplus. As a result, the schemes differ in their risk allocation rules, which specify who of the stakeholders, when, and to what extent is taking part in risk-bearing. In a defined benefit scheme with contribution adjustment (DB_{CA}), the workers bear the funding risks, whereas in a defined benefit scheme with benefit adjustment (DB_{BA}), the retirees bear the funding risks. In the hybrid defined benefit (DB_H) scheme, both workers and retirees bear the funding risks. Ex ante, contributions and benefits are set such that in expectation, any new entry generation funds its own pension. Ex post, a given generation may be a net payer who leaves a surplus for the future generations, but also may be a net receiver who leaves a deficit to the future generations. In this way, funding risks can be shared among many generations during a long period of time. The asset allocation and risk allocation rules are chosen at an initial time and are invariable. The collective pension schemes are optimized with respect to the contribution rates, the adjustment coefficients (i.e., the risk allocation rule) and the portfolio weights, both from the point of view of a newly entering cohort, as well as from the point of view of a social planner who also takes the utility of future generations into account. We also examine the market value of contributions and benefits derived from the optimal schemes, to check whether the schemes are actuarially fair.

Our main findings are the following. (*i*) From the entry cohort point of view, the well designed realistic collective pension schemes can be welfare improving over the optimal individual life-cycle benchmark. In a stylized example, the hybrid collective scheme leads to a 2.3% increase in certainty equivalent consumption per annum vis-à-vis the optimal individual pension scheme. Despite the fact that adjustment speed coefficients and asset allocation are fixed, the hybrid scheme still outperforms the individual benchmark. (*ii*) Due to IRS, participants are more capable to exploit the positive equity premiums by taking more risk. (*iii*) Hybrid pension plan with flexibility in adjusting both contribution and benefit levels to absorb funding residue are the most preferred in welfare terms. It outperforms plans that only allow for adjustments in either contributions or benefit levels. (*iv*) The welfare gains for the new entry cohort are not at the cost of older and future cohorts. From an ex-ante point of view, the expected welfare of the future cohorts is even higher compared to the currently entering cohort. The key intuition is that, the long term nature of the collective pension funds allows for smoothing of risks across many generations, and risk pooling or forming a mutual insurance across generations is welfare enhancing for all from the ex ante perspective. (*v*) Although optimization occurs by the entering generation, the fund is expected to build up a buffer, and is, in expectation, overfunded in the long run. Therefore, both intertemporal risk smoothing strategies (i.e., IRS and the asset accumulations) can be used within combination in collective pension funds to improve intertemporal smoothing.

There are a few differences between our paper and the existing literature. Allen and Gale (1997) and Gollier (2008) study the role of shareholder-sponsored financial intermediary (or corporate-sponsored pension funds) in facilitating intertemporal risk smoothing. They find large welfare improvements if shareholders are able to contribute or accumulate sufficiently large financial buffer as means for intertemporal smoothing. Importantly, they point out that, without mandatory participation, the system will break down and go back to the market solution once the buffer is exhausted or become insufficient. In addition to the potential thread from underfunding, Van Bommel (2007) argues that intergenerational risk sharing will not be sustainable when the fund is overfunded, as new generations might renegotiate and raid the surplus. These studies suggest that without compulsory entry and/or with the opportunity to renegotiating risk sharing rules, funded IRS-pension schemes cannot improve on the market equilibrium in which agents act individually. However, their models are much more abstract, with two or three overlapping generations respectively. Our OLG model is much more realistic, with 55 generations. In such a world, it is realistic to assume that participations can be enforced, and asset-raids can be avoided.

Furthermore, we extend the IRS literature with applications in funded pension systems. Traditionally, intergenerational transfers are implemented by government through fiscal policy and public debt management, monetary policy and Pay-As-You-Go social security programs. It is well known from public finance literature that well designed intergenerational risk sharing under mandatory participation could be welfare improving, especially when dealing with systematic risks. For a general exposition of these issues see Gordon and Varian (1988) and Shiller (1999)). Some specific papers in this area are Fisher (1983), Gale (1990), and Bohn (2003) about fiscal policy and public debt management; Weiss (1979) about monetary policy; Merton (1983), Enders and Lapan (1982), and Krueger and Kubler (2006) about social security programs. In the pension fund context, Teulings and De Vries (2006) argue that it is welfare improving if individuals can borrow against their future labour income and invest before entering the labour market, and expect that pension funds with IRS has similar effect. Instead of assuming borrowing, we model IRS within funded pension schemes with risks borne by multiple overlapping cohorts.

The recognition of the welfare aspects of risk sharing within pension funds is important. An analysis in terms of only market value may easily lead to the spurious conclusion that pension funds are irrelevant, as argued for example by Exley (2004). Indeed, the market value of contributions equals the market value of benefits in our collective pension schemes, and hence a zero-sum game in value-terms; But the schemes with IRS are potentially welfare enhancing, and thus a positive-sum game in welfare terms. These results have important implications related to the social security reforms worldwide. Due to the increasing demographic pressure, many countries are gradually reducing (or planning to reduce) the PAYG social security and promoting the individual defined contribution saving schemes (see for instance recent discussions of Feldstein (2005) and Diamond and Orszag (2005)). Our results show that the individual defined contribution scheme is not the optimal funded scheme, even in a frictionless world with rational and sophisticated individuals. Collective funded schemes with well organized intergenerational risk sharing mechanisms are better choices from a welfare perspective.

The paper is organized as follows: Section 2.2 describes the structure of the economy and the preferences of the pension scheme participants. Section 2.3 introduces collective pension arrangements that allow for intergenerational risk sharing. Section 2.4 analyzes the optimal collective pension schemes and compares that with the optimal individual scheme benchmark, and looks at the implications for old and future cohorts. Section 2.5 discusses the robustness and sustainability of these schemes. Section 2.6 concludes.

2.2 Economy, overlapping generations and preferences

Two asset classes are traded in the financial market: risky stock index and risk-free assets. For simplicity, we assume that the interest rate, r, is non-stochastic. Stock prices follow a geometric Brownian motion with a constant drift μ. The investment portfolio is a mix of stocks and risk-free assets, with portfolio weight

ω_t for stocks. The dynamics of this asset portfolio are given by

$$dW_t/W_t = [r + \omega_t(\mu - r)]dt + \omega_t \sigma dZ_t \tag{2.1}$$

where σ is the volatility of stock returns. Both the expected return and the volatility increase linearly with the fraction ω_t invested in equities. The Sharpe ratio of the investment portfolio equals

$$\lambda = \frac{\mu - r}{\sigma} \tag{2.2}$$

The stochastic discount factor M_t in this economy is the deflator for risky cash flows, which in our model evolves according to

$$dM_t/M_t = -rdt - \lambda dZ_t \tag{2.3}$$

This deflator can be used to calculate the market value of stochastic pension contributions and benefits.

The default values for the model parameters are in line with the values used in the recent literature. We assume the (real) interest rate is constant and equals $r = 2\%$. The expected (real) return on stocks is assumed to be $\mu = 6\%$, implying an equity premium of 4%, which is in line with the long run estimates in Fama and French (2002). The volatility of stock returns is $\sigma = 15\%$, which is close to the value (15.7%) estimated by Cocco, Gomes and Meanhout (2005).

We consider an economy populated by 55 overlapping generations (from age 25 to 80). We assume that all individuals start working at age 25 ($t = 0$), retire at age 65 ($t = R = 40$), and die at age 80 ($t = 55 = T$). During the first 40 years people work and earn a flat real labor income which is normalized to 1.[3] The population (as well as the collective pension fund) has stationary age composition. During the retirement period ($R \leq t < T$), the individuals receive no income but consume their accumulated pension wealth, denoted by W_t. Apart from the pension wealth, there are no savings. The 55 homogeneous cohorts have the same population size and share the same preferences. Individuals have constant relative risk aversion

[3] All variables throughout this paper are expressed in real terms, i.e. scaled by the price level. It is assumed that wage inflation is identical to price inflation.

(CRRA) utility function defined over a single nondurable consumption good. Let c_t denote the consumption level at time t, then the individual's preferences are defined by

$$U = E_0 \left[\int_0^T e^{-\delta t} \frac{c_t^{1-\gamma}}{1-\gamma} dt \right] \tag{2.4}$$

where γ is the risk aversion parameter and δ is the subjective discount rate. In the baseline setup, we set the subjective discount rate equal to $\delta = 4\%$, and the risk aversion parameter $\gamma = 5$ for a typical individual. In the robustness checks we allow for a lower equity premium and different values of δ and γ.

In the later sections, we use certainty equivalent consumption (CEC) as the welfare measure to gauge the performance of collective and individual pension schemes. The CEC be easily backed out from the following equality

$$U = \int_0^T e^{-\delta t} \frac{(CEC)^{1-\gamma}}{1-\gamma} dt \tag{2.5}$$

2.3 Pension schemes with intergenerational risk sharing

In this section, we model collective pension schemes that allow for intergenerational risk sharing. Mandatory participation is required by law. We first define the targeted pension benefit and contribution policies for such funds, and then specify the risk allocation rules which adjust the benefits and/or contributions as functions of finding surplus. Finally the design parameters are optimized with respect to a chosen objective.

2.3.1 Pension liability

Since the collective pension schemes considered in this paper are certain variants of defined benefit (DB) schemes, we shall start with traditional DB schemes to introduce the target liability, target benefit and contribution rates. In a traditional average salary defined benefit scheme, the retirement benefit is a fixed fraction (the so-called replacement rate) of labor income during the working period. These pension benefits are funded by the contributions plus investment proceeds. There is a one-to-one relationship between the replacement rate b and the contribution rate p (recall that we have normalized the flat real labor income to 1). The higher the pension ambition, the more contributions are required. The actuarially fair contribution principle requires that *ex-ante* each generation finances its own pension. That is, the market value of the contributions equals the market value of the benefits. The actuarially fair contribution rate can be solved from the following present value equivalence:

$$\int_0^R e^{-rs} p ds = \int_R^T e^{-rs} b ds. \tag{2.6}$$

As described in Section 2.2, 55 homogeneous overlapping generations (40 working and 15 retired cohorts) coexist in the pension fund at any point in time. At the aggregate level, the sum of benefits over all retirees is $15b$, and the sum of contributions over all workers is $40p$.

For each age cohort x, the target DB liability equals the difference between the present value of risk-free benefits minus the present value of yet-to-be-paid risk-free contributions. The liabilities of the fund can be calculated simply as the sum of the liabilities for each age cohort:

$$L = \int_0^R \left(\int_R^T e^{-r(t-x)} b dt - \int_x^R e^{-r(t-x)} p dt \right) dx + \int_R^T \left(\int_x^T e^{-r(t-x)} b dt \right) dx \tag{2.7}$$

Given the stationary age composition of the fund and the fixed target benefit level, the target liability L is time-invariant. If the fund invests full in risk free asset, then the actual liability follows exactly as the target liability L. If the fund accepts mismatch risk, for example by investing in stocks, there may be funding surpluses or deficits with respect to the target liability. Let A_t denote the value of the pension fund's assets, which starts off with an initial asset value of $A_0 = FR_0 \cdot L$, where FR_0 is the initial funding ratio. Let ω denote the fraction of assets invested in risky assets. The portfolio weight ω is time-invariant and the same for each cohort in the fund. This assumption is motivated by the observed common practice in collective pension funds, where the same investment policy is implemented for all cohorts. Then, the fund's assets follow the dynamics

$$dA_t = [A_t (r + \omega (\mu - r)) + 40p_t - 15b_t] dt + \omega \sigma A_t dZ_t \tag{2.8}$$

where p_t and b_t denote the actual contribution and benefit levels which will be specified shortly. In most of the experiments we assume $FR_0 = 1$, but we shall show some results with initial over- or under-funding. The fund surplus is defined as the difference between assets and the target liability level:

$$S_t = A_t - L \tag{2.9}$$

2.3.2 Risk allocation rules

If the investments of the fund were completely risk free, there would never be any overfunding or underfunding. However, with a risky investment policy, or other sources of systematic risks, the mismatch risk has to be allocated in some way. This is done, for instance, by adjusting contributions p_t and benefits b_t as a function of the fund surplus S_t and the target contribution (p) and benefit (b). This adjustment scheme is motivated by real-world examples.[4] We now discuss three stylized schemes, that differ in their risk allocation rules. These risk allocation rules specify who of the stakeholders, when, and to what extent is taking part in risk-bearing.

In the defined benefit with contribution adjustments (DB_{CA}) scheme, benefits are fixed at $b_t = b$, but contributions are adjustable. So, the working cohorts bear all the funding risk. We specify a simple contribution policy, where contributions per cohort are a function of the target contribution level p and the funding residual per cohort S_t/R :

$$p_t = p - \alpha S_t/R \tag{2.10}$$

The slope coefficient α determines the speed of absorbing the funding imbalances. The choice $\alpha = 1$ implies that a funding imbalance is immediately fully absorbed. A lower value of α implies that part of the funding residual is shifted to the future, and shared across generations. Therefore the lower the value of α, the higher the

[4] We refer to Ponds and Van Riel (2007) for a more detailed description.

degree of intergenerational risk sharing.

In the defined benefit with benefit adjustments (DB_{BA}) scheme, contributions are fixed at $p_t = p$, but benefits are adjustable in order to absorb the funding surplus. In this scheme, retired cohorts bear the funding risk, where benefit per retired cohort are a function of the target benefit level b and the funding residual per cohort $S_t/(T-R)$:

$$b_t = b + \beta S_t/(T-R) \tag{2.11}$$

Again, the lower the value of β, the higher the degree of intergenerational risk sharing.

The hybrid defined benefit (DB_H) scheme adjusts both contributions and benefits simultaneously to absorb the funding residual. The contribution and benefit are linearly related to the funding residual S_t. More specifically, a fraction α of the funding residual is shared among employees and a fraction β of the funding residual is shared among retirees:

$$p_t = p - \alpha S_t/R \tag{2.12}$$

$$b_t = b + \beta S_t/(T-R) \tag{2.13}$$

Under this rule, consumption before and after retirement is related to the funding surplus. This resembles the optimal consumption rule in the classical consumption and portfolio choice problem in the spirit of Merton (1969) where the optimal consumption is linearly related to wealth.

The following table summarizes the differences among the collective schemes:

Defined benefit with contribution adjustments (DB_{CA})	$\alpha > 0$	$\beta = 0$
Defined benefit with benefit adjustments (DB_{BA})	$\alpha = 0$	$\beta > 0$
Hybrid defined benefit (DB_H)	$\alpha > 0$	$\beta > 0$

When $\alpha + \beta$ approaches the risk free rate r, each generation in each period absorbs only the 'interest' accrued on their funding residual and passes the principal into the infinite future.

2.3.3 Pension policies optimization

In practice the board of trustees decides upon the pension policies, taking into account the preference of the participants. In line with this practice, we assume the board of trustees optimally chooses the target contribution rate (p) together with the risk allocation rule (α, β) and investment policy (ω), based on the expected life-time utility of the 25-year-old entering participants. That is, the policy parameters $\{p, \alpha, \beta, \omega\}$ of the three collective plans are optimized from the perspective of the cohort entering at time $t = 0$. Therefore, the specification of the utility function of the pension fund is the same as in the individual case (2.15). The optimization problem of the pension fund then becomes

$$\begin{aligned} U &= \max_{\{p,\alpha,\beta,\omega\}} E_0 \left[\int_0^T e^{-\delta t} \frac{c_t^{1-\gamma}}{1-\gamma} dt \right] \qquad (2.14) \\ \text{with} \quad c_t &= 1 - (p - \alpha S_t / R), \quad t < R \\ c_t &= b + \beta S_t / (T - R), \quad R \leq t \leq T \end{aligned}$$

subject to the wealth dynamics (2.8). Since the pension fund is borrowing constrained as well as short-sale constrained, the portfolio choice is constrained by $0 \leq \omega \leq 1$. The actual contributions and benefits (p_t, b_t) follow directly from the risk allocation rules (2.12, 2.13) in different schemes. To get stability over time (i.e., non-explosive values for surpluses), we need to restrict $\alpha + \beta \geq r$. The decisions $\{p, \alpha, \beta, \omega\}$ are made at an initial time and remain time invariant. The optimization problem therefore is static and can be solved using Monte Carlo simulations combined with a standard grid search in four dimensions. Appendix A gives more details about the optimization procedure.

2.4 Evaluation of pension schemes

In this section we evaluate the performance of the optimal collective schemes from welfare perspective. First, in subsection 2.4.1, we introduce the benchmark, namely the optimal individual scheme benchmark. Subsection 2.4.2 then compares the optimal collective schemes with the individual benchmark from the entry cohort's perspective. Subsection 2.4.3 evaluates the optimal schemes from older and future cohorts' perspective. Subsection 2.4.4 discusses additional sensitivity analysis.

2.4.1 Optimal individual scheme benchmark

We consider two individual schemes, where the individual bears all the non-diversifiable investment risks and there is no intergenerational risk sharing. The first scheme is the optimal individual scheme, where the individual can optimally choose his consumption level and portfolio composition (under short sales and borrowing constraints) at any time throughout his life. This scheme serves as the benchmark for the collective schemes with IRS, which have been presented in Section 2.3. The second scheme is a defined contribution scheme with a fixed contribution rate which is chosen optimally at the beginning of the working life, but the consumer is still able to adjust the portfolio weight throughout his life. The later scheme captures common practice in the real world and is more realistic than the optimal individual scheme; Poterba, Rauh, Venti and Wise (2005) use a similar setup.

The optimal individual scheme (OI) implements the optimal life cycle consumption and portfolio choice of an individual investor under borrowing and short-selling constraints. During working period, the investor receives his labor income, which is normalized to 1. The individual chooses his consumption stream c_t optimally at every point in time, and invest the rest in financial markets. After retirement, the individual optimally consumes down his wealth. The individual is free to choose the portfolio weight of stocks ω_t continuously, within the short-selling and borrowing constraints, i.e. $0 \leq \omega_t \leq 1$. Formally, the individual's consumption and portfolio choice problem is characterized by the following objective function

$$U = \max_{\{0 \leq \omega_t \leq 1, c_t\}} E_0 \left[\int_0^T e^{-\delta t} \frac{c_t^{1-\gamma}}{1-\gamma} dt \right] \tag{2.15}$$

and subject to the pension wealth dynamics

$$dW_t^{OI} = [W_t^{OI}(r + \omega_t(\mu - r)) + 1 - c_t]dt + \sigma\omega_t W_t^{OI} dZ_t \quad (0 < t < R) \tag{2.16}$$
$$dW_t^{OI} = [W_t^{OI}(r + \omega_t(\mu - r)) - c_t]dt + \sigma\omega_t W_t^{OI} dZ_t \quad (R \leq t < T) \tag{2.17}$$
$$W_0^{OI} = 0 \tag{2.18}$$

where W_t^{OI} is the accumulated financial wealth in the OI scheme.

Because of the short-selling and borrowing constraints, no analytical solution is available. We use dynamic programming to solve for the consumption and investment in the optimal individual scheme. We use the numerical procedures presented by Carroll (2006) to solve for the optimal consumption and portfolio policy before retirement, under the borrowing and short-selling constraints, $0 \leq \omega_t \leq 1$. Details of the solution method are given in Appendix B.1.

The defined contribution (DC) scheme differs from the optimal individual scheme in that the individual contributes a fixed fraction of his labor income into the DC scheme and consumes the rest of his income. At the entry date, the contribution rate m is fixed for the entire the working period ($0 < m < 1$ during the working period and $m = 0$ after retirement). The individual's consumption before retirement is therefore $c_t = (1 - m)$, and after retirement the consumption is optimally chosen. The individual adjusts his portfolio (ω_t) throughout life. Formally, the individual's consumption and portfolio choice problem is characterized by the following preferences

$$U = \max_{\{m, 0 \leq \omega_t \leq 1, c_t\}} E_0 \left[\int_0^R e^{-\delta t} \frac{(1-m)^{1-\gamma}}{1-\gamma} dt + \int_R^T e^{-\delta t} \frac{c_t^{1-\gamma}}{1-\gamma} dt \right] \tag{2.19}$$

subject to the pension wealth dynamics

$$dW_t^{DC} = [W_t^{DC}(r + \omega_t(\mu - r)) + m]dt + \sigma\omega_t W_t^{DC} dZ_t \quad (0 < t < R) \tag{2.20}$$
$$dW_t^{DC} = [W_t^{DC}(r + \omega_t(\mu - r)) - c_t]dt + \sigma\omega_t W_t^{DC} dZ_t \quad (R \leq t < T) \tag{2.21}$$
$$W_0^{DC} = 0 \tag{2.22}$$

The optimal contribution rate m is found by numerical grid search, detailed in

Appendix B.2. Since m and p both refer to a fixed contribution rate, the notation p is used in the tables for both collective and individual schemes.

Figure 1 shows the life cycle quantile distributions (at 5%, 50% and 95% quantile levels) of portfolio weight, consumption and wealth accumulation of the two individual schemes for a benchmark investor ($\gamma = 5$, $\delta = 4\%$). The left panel shows the results for the OI scheme and the right panel shows the DC scheme. The distribution of portfolio choices (ω_t) is shown in the top panel of Figure 1. On average, the optimal portfolio weights are decreasing with age. As explained by Campbell and Viceira (2002), this is due to the leverage effect of human capital (or to be precise, the present value of future pension contributions). The desired equity exposure in total wealth, including financial wealth and human capital, is constant over the life cycle. Individuals achieve the desired equity exposure of their total wealth by adjusting their financial portfolio.

The distribution of consumption (normalized by income) is shown in the middle panel of Figure 1. The median consumption level is relatively flat over the life cycle, but uncertainty in old-age consumption is increasing with time. In the OI scheme, the consumption is higher (hence new saving is lower) when the financial wealth is boosted by good returns. Finally, the accumulated assets over life cycle are presented in the lower panel of Figure 1. These asset values show wide dispersion and peak at the retirement date. The median amount saved at retirement is around 11 times of the annual labor income for OI scheme or 13 times for DC scheme. The above results are based on baseline parameter assumptions. Further sensitivity analyses with respect to risk aversion and subjective discount rate can be found in Appendix B.3.

2.4.2 Optimal collective schemes and welfare evaluation

Having described both collectives and individual schemes in previous sections, Table 2.1 reports the optimal design parameters $\{p, \alpha, \beta, \omega\}$ and the obtained welfare levels for those schemes under the default parameter values ($r = 2\%$, $\delta = 4\%$, $\mu - r = 4\%$, $\gamma = 5$). It also shows the results under two alternative risk aversion assumptions ($\gamma = 3$ and $\gamma = 8$). All schemes are initially fully funded (i.e. $A_0 = L$).

We find that institutional settings have a significant impact on the optimal contribution rates and investment policies. The target level of contribution p and portfolio choice ω vary significantly across schemes. The optimal contribution rate in DC scheme is the lowest of all, which is 11%. The optimal target contribution rate in DB_{CA} scheme is the highest, requiring 16.6% of annual salary. This leads to the highest risk-free replacement rate of 78.4% after retirement. The desired contribution rate in DB_{BA} scheme is the lowest among the collective schemes, requiring 13.1% of salary, which is close to the DC scheme. Consequently, its target replacement rate b is also the lowest at 53.6%. The hybrid scheme DB_H requires 14% of salary as contribution rate and leads to a 66% target replacement rate. The actual contributions and benefits are different from these risk-free levels, as the pension fund can run surpluses or deficits.

The portfolio choice of DB_{BA} scheme is less aggressive than the other collective schemes. The risky portfolio weight of DB_{BA} scheme is 60%, whereas the risky share of DB_{CA} scheme reaches 96%. The hybrid scheme DB_H accommodates the most aggressive investment portfolio, namely 100% in stocks. Comparing with the portfolio choices in the optimal individual schemes (Figure 1), where the stock portfolio share declines to 35% over the life cycle, the stock allocations of collective schemes are much higher. This shows that intergenerational risk sharing makes an aggressive investment policy more attractive. Due to intergenerational risk sharing, participants are more capable to exploit the positive equity premiums. This finding confirms the results of Gollier (2008) that intergenerational risk sharing increases the demand for risky investment. The values of the contribution and benefit adjustment parameters α and β are typically small, indicating that funding mismatches are absorbed gradually over time.

Figure 3 provides some graphical insight in the distribution of consumption over time. The figure plots the 5%, 50% and 95% quantiles of the normalized consumption, c_t, of the optimal collective plans, with the scheme parameters set corresponding to Table 2.1 with $\gamma = 5$. In general, the consumption profiles are upward sloping with age. The distributions of consumption indicate that contribution reductions are more frequent, and also larger, than contribution increases. This is due to the positive equity premium. The figure shows that the hybrid scheme (DB_H) is much better in distributing shocks over the full lifetime of the individual than the other collective schemes, where there are large fluctuations in consumption either just before retirement (DB_{CA}) or during the retirement period (DB_{BA}). This demonstrates that more efficient risk sharing can be achieved by using more risk absorbers. Through adjusting *both* contribution and benefit policies, the DB_H scheme spreads the funding residual over both workers and retirees, i.e. all cohorts are involved in the risk sharing. From the risk allocation perspective, DB_{CA} and DB_{BA} represent two extremes, with one group of people bearing all risks, while the other group is fully insured. The hybrid scheme allocates the risks more efficiently among all generations, hence reducing the costs of risk taking and resulting in welfare improvements.[5]

The most important result in Table 2.1 is the welfare levels of the optimized pension schemes from the entry cohort perspective. Welfare level is reported as the normalized certainty equivalent consumption in units of annual salary, CEC. Let CEC^{OI} denote the welfare level achieved by the optimal individual scheme (OI). The ratio CEC/CEC^{OI} then shows the welfare of the collective schemes relative to this optimal individual benchmark. For participants with a default risk aversion ($\gamma = 5$), the hybrid defined benefit (DB_H) scheme provides a welfare gain of 2.3% per annum vis-a-vis the OI scheme in terms of certainty equivalent consumption. Over the full life-cycle, this amounts to more than one annual salary gain. The results indicate that well-structured intergenerational risk sharing is welfare enhancing compared to the optimal individual scheme. Despite the fact that adjustment speed coefficients and asset allocation are fixed, the hybrid collective scheme still outperforms the individual benchmark. The hybrid pension plan with flexibility in adjusting both contribution and benefit levels to absorb funding residue are the most preferred in welfare terms. It outperforms other collective plans that only allow for adjustments in either contributions or benefit levels. Particularly, the DB_{CA} scheme has a welfare loss of 0.3% relative to the optimal individual scheme, whereas the DB_{BA} scheme shows a welfare loss of 2.7%. Notice that the welfare of the DB_{BA} scheme is very close to the welfare of its individual counterpart, the DC plan. As the DC scheme induces a 2.8% welfare

[5] This finding confirms the results of Van Hemert (2005), who shows that the optimal intergenerational risk sharing in social security program is neither a pure defined benefit type nor a defined contribution type, but a state contingent hybrid scheme.

loss vis-a-vis the OI scheme, the hybrid defined benefit scheme outperforms the more realistic DC scheme by 3–6% per annum.

Furthermore, in Section 2.5.1, we will show that the market value of contributions equals the market value of benefits in the optimal collective schemes. It means that the contributions and benefits set according to Equation (2.6) and the risk sharing rules in Equations (2.12) and (2.13) are actuarially fair. Ex ante, the starting generation does not borrow from future generations and therefore there is no "chocolate paradox" of the type described by Shell (1971). More details of the discussion is postponed till Section 2.5.1.

2.4.3 Welfare of older and future generations

We have shown in the previous subsection that the entry cohort is better off by starting with an optimal DB_H scheme at time 0. What about the welfare of the coexisting older generations (from age 26 to age 78)? Are they better off or worse off by switching to the optimal DB_H scheme as compared with switching to OI schemes at time 0 with accumulated assets equal to the value of their accumulated liabilities? Figure 4 compares the welfare levels of the living older generations under the optimal DB_H scheme, with what would be achieved under the OI scheme if the older generations switch to their own OI schemes, $W^{OI}_{x,t=0} = L_x$. The welfare of older generations is defined as the CEC of x-year-old under the given optimal schemes governed by $\{p^*, \alpha^*, \beta^*, \omega^*\}$:

$$\int_{x-26}^{79-x} e^{-\delta t} \frac{(CEC_x)^{1-\gamma}}{1-\gamma} dt = E_0 \left[\int_{x-26}^{79-x} e^{-\delta t} \frac{c_{x,t}^{1-\gamma}}{1-\gamma} dt \right], \quad \text{for } x\text{=26,...,78} \tag{2.23}$$

Figure 4 shows that the welfare levels of the older generations are higher than (or indistinguishable from) what would be obtained under the OI scheme, if they had switched at time 0. Therefore the welfare gain the entry cohort does not cause a welfare loss for the living older generations.

The question whether or not the welfare gain of the current new-entry cohort comes at the expense of the future cohorts is an important issue for pension plan design. To investigate this question, we compute the ex ante expected welfare of the future cohorts, who will join the optimal collective schemes described as in

Table 2.1, with the initial condition that the pension funds are fully funded at $t = 0$. We calculate the certainty equivalent consumption, CEC^f, of the future cohorts (the future 25-year-old's) who will enter into the pension fund in f years' time, with $f = 0, 1, 2, ..., 1000$ years. The expected welfare of cohort f obtained from participating a given pension scheme, during his life time from year f till $f + T$, is measured in terms of certainty equivalent consumption as follows:

$$E_0 \left[\int_0^T e^{-\delta x} u \left(c_{f+x}^f \right) dx \right] \equiv \int_0^T e^{-\delta x} u \left(CEC^f \right) dx \tag{2.24}$$

Table 2.2 presents the CEC of the future cohorts (up to cohort $f = 1000$) relative to that of the optimal individual scheme, CEC^f / CEC^{OI}, which can be seen as the relative welfare gain or loss over the individual benchmark. The expected welfare of all future cohorts are higher than the time 0 entry cohort. Figure 5 (upper panel) visualizes the welfare improvements including the 1000 future cohort. The welfare is steadily increasing for future cohorts and converging to a level much above the level for the entry cohort.

From this result we can conclude that the expected welfare gain of the current entry cohort is not at the cost of the future cohorts from an ex ante perspective. Although optimization occurs by the entering generation, the fund is expected to build up a buffer, and is, in expectation, over-funded in the long run, as shown in Figure 5 (lower panel). Due to these transfers, future cohorts are able to benefit from the positive equity premium before they are born. Intergenerational risk sharing together with asset accumulations greatly enhance the intertemporal risk smoothing capability of the economy with collective pension schemes.

2.4.4 Sensitivity analysis

We now perform some robustness checks to see if our results still hold under different risk and time preferences, and with a reduced equity premium. The degree of risk aversion has several effects on the optimal pension design, as shown in Table 2.1. Comparing the results as the coefficient of risk aversion increases from $\gamma = 3$ to $\gamma = 8$, one observes three changes: (*i*) the portfolios become less risky, (*ii*) the contribution rates increase, and (*iii*) the schemes rely more on intergenerational risk sharing by choosing lower values for α and/or β. Furthermore, the degree of risk aversion affects the welfare gain from intergenerational risk sharing. The welfare gain for a less risk averse investor ($\gamma = 3$) is 3.9% in the DB_H scheme, whereas it is reduced to 0.9% for a more risk averse investor ($\gamma = 8$). The result that more risk-averse agent obtains less welfare gain may seem counter-intuitive at first sight, but observe that the less risk-averse agent is willing to accept a more risky portfolio and thus benefits more from intergenerational risk sharing. The less risk averse participants are also willing to accept larger adjustments in contributions and benefits in order to absorb funding mismatches, which is reflected in higher values for α and β.

Table 2.3 shows the optimal schemes and the welfare evaluations when the equity premium is reduced from 4% to 3%. The results are in line with our baseline results reported in Table 2.1, supporting our claim that well-organized IRS by collective pension funds is welfare improving. However the welfare levels are lower and the relative welfare gains are smaller. For instance, the welfare gain for participants in DB_H scheme vis-à-vis the OI scheme ranges now from 0% to 2%, depending on the degree of risk aversion, and 2% to 4% vis-à-vis the DC scheme. A lower subjective discount factor results in higher contribution rates, similar to the individual schemes. However, other design parameters are not significantly different from the baseline cases.

2.5 Sustainability

Having shown that IRS in collective pension plans can be welfare improving with mandatory participation and renegotiation-proof, we investigate in this section several issues related to the sustainability of the pension schemes. Sustainability may be at stake when new cohorts are reluctant to step into the pension plan because of underfunding. In this section we therefore consider the market value of the pension deals, the distribution of the funding ratio and possible social planner's designs (which takes future generations into account).

2.5.1 Market valuation of the pension schemes

In this section, we focus on the market values of contributions and benefits in the optimal collective schemes. For an individual member, it may be difficult to trade his pension claims once he stepped into a pension contract, but a market valuation is useful ex-ante to evaluate a given pension deal. Furthermore, it is important to investigate whether the target contributions and benefits set according to Equation (2.6) and the risk sharing rules in Equations (2.12) and (2.13) are actuarially fair. If so, then it means the market value of contributions equals the market value of benefits, and ex ante, the starting generation does not borrow from future generations and therefore there is no "chocolate paradox" of the type described by Shell (1971).

From the perspective of a new pension fund member, the market value of the pension deal is the value of the actual benefits minus the contributions. We take all the (stochastic) cash flows (p_t, b_t) generated by the optimal pension schemes and value them using the deflator method.[6] For the typical member with retirement date R and life expectancy T, the net present value (NPV) of the pension deal is the difference between the present value of benefits (PVB) and the present value of contributions (PVP)

$$NPV = PVB - PVP = E_0 \left[\int_R^T M_t b_t dt \right] - E_0 \left[\int_0^R M_t p_t dt \right] \tag{2.25}$$

where p_t and b_t are the (stochastic) actual contributions paid and benefits received. From an ex-ante point of view, the pension scheme is a fair deal if the NPV is zero. Figure 2 shows quantile plots of the normalized funding ratio's, $FR_t = S_t/L$, of the three optimal collective schemes. Funding surpluses are more frequent (and larger in size) than funding deficits. This is due to the positive equity premium. However, this does not add value to the pension deal: the lower contributions occur in scenarios where the equity returns are high, but the deflators in such scenarios are low. Therefore, the present values of the paid contributions and the cost-price contributions are equal.

[6] This analysis builds on Bader and Gold (2002), Blake (1998), Chapman et al. (2001) and De Jong (2007), and results in the method of value-based generational accounting of Ponds (2003).

The collective schemes may shift the funding mismatches beyond one's life time, leaving surplus or deficit in the notional cohort account. An alternative way of calculating the present value of pension deals therefore is to look at the remaining balance left in the notional cohort account at the end of life of a cohort. A positive remaining balance indicates positive transfers from this generation to other cohorts. Similarly, a negative transfer means the cohort receives cash flows from other cohorts. In a collective pension scheme, any cohort writes a call option to share a funding surplus with the other cohorts, and holds a protective put option to receive protections from the other cohorts. When the value of the call equals the value of the put, the pension deal is a fair deal in value terms ex ante. Formally, let a_t denote the accumulated asset in the notional cohort account through out their life time $(0 \leq t \leq T)$. Each cohort starts its notional account with no surplus or deficit, i.e. $a_0 = 0$. The dynamics of the cohort account are

$$da_t = \left[a_t \left(r + \omega \left(\mu - r\right)\right) + p_t - b_t\right] dt + \omega\sigma a_t dZ_t \tag{2.26}$$

Integrating out this dynamics and applying Equation (2.25), the NPV of the pension deal can be decomposed as the sum of the value of the call and the put option, valuing the positive and negative transfers to other cohorts:

$$NPV = -E_0 \left[M_T a_T\right] = -E_0 \left[M_T \left(a_T\right)^+\right] + E_0 \left[M_T \left(-a_T\right)^+\right] = -call + put \tag{2.27}$$

Table 2.4 summarizes the market valuation of the intergenerational transfers and options for initially fully funded pension schemes ($FR_0 = 1$). As required by the assumption of actuarial fairness, the NPV is indeed zero for all collective schemes. However, the market values of positive and negative transfers (the implicit call and put options) are potentially large, between 0.2 to 1.25 annual salaries, indicating a substantial amount of ex post transfers. The magnitude of the transfers depend on the chosen asset mix, the level of contributions and the risk sharing rules. The DB_{CA} scheme results in the highest value of transfers whereas the DB_{BA} scheme results in the lowest value of transfers. This makes the DB_{CA} system less sustainable than the DB_{BA} system, as future generations can be confronted with larger deficits in the pension fund.

2.5.2 Initially underfunded and overfunded schemes

The distributions of the funding status over time are critical for the welfare of the future cohorts and the sustainability of the schemes. Seriously underfunded situations are difficult for the sustainability of the fund, as in these situations future generations may want to step out.

Table 2.5 shows the welfare gains and losses of entering the initially over- and underfunded schemes (with funding ratios being $FR_0 = 1.1$, $FR_0 = 0.9$ and $FR_0 = 0.8$ respectively). These schemes are set at their optimal designs as in Table 2.1 (middle panel). The new cohort joining an underfunded collective scheme is not necessarily worse off in welfare terms, comparing with the optimal individual benchmark. For instance, when FR_0 falls to 0.9 (0.8) initially, the welfare of DB_H is 100.7% (98.8%) of that of the optimal individual scheme. Hence it is possible for well-structured collective pension schemes to absorb funding deficits up to 10% to 20% by intergenerational risk sharing and still enhance the welfare for her participants.

Table 2.6 shows the market valuation of the intergenerational transfers when the collective schemes are initially underfunded. The net transfers ex ante are non-zero. The net transfers are large and positive, meaning the entry cohort makes up a large part of the initial deficits, by either higher contributions (DB_{CA}, DB_H) or lower benefits (DB_{BA}, DB_H). Furthermore, the net transfers are proportional to the degree of underfunding. For instance, the value of net transfers doubles when FR_0 is reduced from 0.9 to 0.8. Initially overfunded schemes are also reported in Table 2.5 and 2.6. The market values of the net transfers starting with $FR_0 = 1.1$ are mirror images of $FR_0 = 0.9$.

2.5.3 Social planner's view

An alternative way to design the collective schemes is to take into account the welfare of future generations by taking a social planner's view. Suppose the social planner's objective is to maximize a weighted sum of entry and future cohorts' life time expected utility by optimizing the pension scheme $\{p, \alpha, \beta, \omega\}$ at time zero:

$$U^{Social} = \max_{\{0 \leq \omega \leq 1, p, \alpha, \beta\}} E_0 \left[\sum_{f=0}^{\infty} \left(\Delta^f \int_0^T e^{-\delta x} u \left(c_{f+x}^f \right) dx \right) \right] \tag{2.28}$$

where $0 < \Delta < 1$ is the social planner's weighting factor for the future cohorts. There is no clear guidance as to how Δ should be chosen. Gollier (2008) imposes that the assets of the pension fund follow a martingale; this assumption implies $\Delta = 0.937$, given his other parameter values. We assume the weighting factor of the social planner is the same as individuals' subjective discount factor, hence $\Delta = e^{-\delta} = 0.96$.[7] We calculate the social planner's designs for individuals with risk aversion of $\gamma = 5$.

Table 2.7 shows the results for the social planner's optimal pension schemes. Comparing the socially optimal schemes with the schemes optimized for the $f = 0$ entry cohort in Table 2.1, we see that the values of α and β are slightly lower and the portfolio choice becomes more risky. These results indicate that shocks in funding surplus are shared by more generations under the socially optimal schemes. Table 2.8 shows the market value of the intergenerational transfers under the socially optimal schemes. Indeed the market value of transfers (calls and puts) are larger comparing with that of Table 2.4. CEC^{opt} shows the certainty equivalent consumption of the time 0 entry cohort under the scheme optimized by the social planner. From comparing CEC^{opt} with CEC in Table 2.1, we find that the welfare of the time 0 entry generation is slightly reduced under the social planner's schemes. Figure 6 compares the welfare improvements of the DB_H scheme under the socially optimal design versus the $f = 0$ cohort optimal design. In the short run, the entry cohort optimal scheme dominates, but for cohorts entering after $f = 15$ approximately, the socially optimal scheme is better.

These results do not materially change for different risk and time preferences or a lower equity premium; the long run welfare gains for the future cohorts are still substantial.

[7] Feldstein (2005) provides an argument why a 4% discount rate is a reasonable value.

2.6 Conclusion

We have used the institutional setting of funded pension schemes to study welfare aspects of intergenerational risk sharing. Typical for such collective pension plans is that pension benefits and/or pension contributions may depend on the funding status. From the perspective of a newly entering cohort, we optimize the explicit asset allocation and risk allocation rules, which specify who of the stakeholders, when, and to what extent is taking part in risk bearing.

We show that well-designed funded schemes with intergenerational risk sharing are welfare improving over and above the fully optimal individual scheme by up to 2-4% in terms of certainty equivalent consumption. The hybrid defined benefit scheme (DB_H), where risks are shared between working and retired cohorts, performs better than schemes where risks are only borne by workers (DB_{CA}) or retirees (DB_{BA}). The initially fully funded collective pension schemes with intergenerational risk sharing are a zero-sum game in value terms, however they are potentially a positive-sum game in welfare terms. Furthermore, the expected welfare gain of the current entry cohort is not at the cost of the older and future cohorts, from an ex ante perspective. This result has important implications for the current trend of shifting from defined benefit to (individual) defined contribution schemes. In such individual schemes, the risks are concentrated mainly in the retirement phase, giving up the intergenerational risk sharing potential and the greater intertemporal risk smoothing capacity of collective plans.

Introducing more sources of systematic risks in addition to the modeled investment risks, like labor income risk and real interest rate risk, might further strengthen the welfare-enhancing potential of intergenerational risk sharing via collective schemes compared to the optimal individual scheme.

2.7 Appendix A: Solution method of collective schemes

The optimization problem of collective schemes, Equation (2.14), is a static problem, since all the decision parameters $\{p, \alpha, \beta, \omega\}$ are constants. The collective schemes are solved using Monte Carlo simulations and grid search. We first construct a four dimensional grid for the decision parameters. For each combination of the parameters, we simulate the wealth dynamics (2.8) at each point in time, evaluate the resulting funding surplus and determine the state contingent contributions and benefits p_t and b_t according to the risk sharing rules in Equations (2.12) to (2.13). When p_t and b_t are determined, we evaluate the objective function (2.14) for the new entry cohort. A numerical grid search identifies the global maximum of the welfare function (2.14) and its corresponding parameter values $\{p, \alpha, \beta, \omega\}$.

2.8 Appendix B: Solution method of individual schemes

2.8.1 OI scheme

This appendix explains the solution procedure solving the optimal individual (OI) consumption and portfolio choice described in Section 2.4.1. We first rewrite the optimization problem (2.15) in discrete time form (at annual frequency), and rewrite the original objective function in the recursive form as follows

$$
\begin{aligned}
U(W_t) &= u(c_t) + E_t\left[e^{-\delta} U(W_{t+1})\right] && (2.29)\\
s.t.\ W_{t+1} &= (W_t - c_t)\left[R^f + \omega_t\left(\widetilde{R}^s_{t+1} - R^f\right)\right] + 1 \quad (1 \leq t < R) && (2.30)\\
W_{t+1} &= (W_t - c_t)\left[R^f + \omega_t\left(\widetilde{R}^s_{t+1} - R^f\right)\right] \quad (R \leq t < T) && (2.31)
\end{aligned}
$$

where $W_1 = 1$, $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$, $\widetilde{R}^s_{t+1} = \exp\left(\mu + \sigma \int_t^{t+1} dZ_s\right)$, and $R^f = \exp(r)$. The problem is solved using dynamic programming principle, solving backwards from the last to the first period.

For the final period, the optimal consumption $c_T^* = W_T$, and the functional form of indirect utility is known, as $U(W_T) = u(c_T^*) = \frac{(c_T^*)^{1-\gamma}}{1-\gamma}$. Then, we proceed backward in time to $t = T-1$. The procedure for $T-1$ starts by defining a new variable $a_t = W_t - c_t$, as the after-consumption-wealth. Following Carroll (2006), we construct a grid of $a_t = \{a_j\}_{j=1}^J$. Now we solve for the optimal consumption and portfolio policies for each given a_j. The first order conditions with respect to

ω_t and c_t are

$$
\begin{aligned}
0 &= E_t\left[e^{-\delta}U'\left(W_{t+1}\right)\left(\widetilde{R}^s_{t+1}-R^f\right)\right] && (2.32)\\
u'\left(c_t\right) &= E_t\left[e^{-\delta}U'\left(W_{t+1}\right)\left(R^f+\omega_t\left(\widetilde{R}^s_{t+1}-R^f\right)\right)\right] && (2.33)
\end{aligned}
$$

The envelope theorem implies that $u'\left(c_t\right)=U'\left(W_t\right)$, since

$$
U'\left(W_t\right)=E_t\left[e^{-\delta}U'\left(W_{t+1}\right)\left(R^f+\omega_t\left(\widetilde{R}^s_{t+1}-R^f\right)\right)\right] \qquad (2.34)
$$

Replace $U'\left(W_{t+1}\right)$ by $u'\left(c_{t+1}\right)$ in the two first order conditions, we have

$$
\begin{aligned}
0 &= e^{-\delta}E_t\left[\ u'\left(c^*_{t+1}[W_{t+1}]\right)\left(\widetilde{R}^s_{t+1}-R^f\right)\right] && (2.35)\\
c^*_t\left(a_t\right) &= I_u\left(\ e^{-\delta}E_t\left[\ u'\left(c^*_{t+1}[W_{t+1}]\right)\left(R^f+\omega^*_t\left(a_t\right)\left(\widetilde{R}^s_{t+1}-R^f\right)\right)\right]\right) && (2.36)
\end{aligned}
$$

where $c^*_{t+1}[W_{t+1}]$ as the optimal consumption policy at time $t+1$, and $I_u\left(\centerdot\right)$ denotes the inverse function of $u'\left(c_t\right)$. Using any numerical solver, Equation (3.31) will give the optimal portfolio weight $\omega^*_t\left(a_t\right)$ for any given amount of investment a_t. Because of the borrowing and short-selling constraints, we then impose the restriction $0\leq\omega^*_t\leq 1$. Then, Equation (3.32) gives the corresponding consumption $c^*_t\left(a_t\right)$ for any given amount of investment a_t. Finally, the optimal wealth process is endogenously determined by $W^*_t=c^*_t\left(a_t\right)+a_t$. The advantage of this method is that the numerical search is only needed once in solving $\omega^*_t\left(a_t\right)$, while $c^*_t\left(a_t\right)$ can be directly obtained from Equation (3.32).

Following the same procedure illustrated above, we can solve the model backward in time to $t=1$. To generate the average pattern of life cycle portfolio holding we simulate the model from time 1 to T for 10,000 scenario's, and take the average over all simulated scenario's.

2.8.2 DC scheme

The optimal consumption and portfolio choice after retirement in the *DC* scheme is identical to the one in the *OI* scheme. Before retirement, the optimal portfolio weight ω_t is stochastic, depending on the relative size of financial capital and human capital. Campbell and Viceira (2002) show that the optimal individual portfolio choice before retirement in the DC plan without constraints is the leveraged myopic portfolio:

$$\omega_t^* = \frac{1}{\gamma}\frac{\mu - r}{\sigma^2}\frac{W_t^{DC} + PVC_t}{W_t^{DC}} \tag{2.37}$$

where PVC_t denotes the present value of the future pension contributions, as a fraction of the investor's human capital,

$$PVC_t = m\int_t^R e^{-rs} y ds \tag{2.38}$$

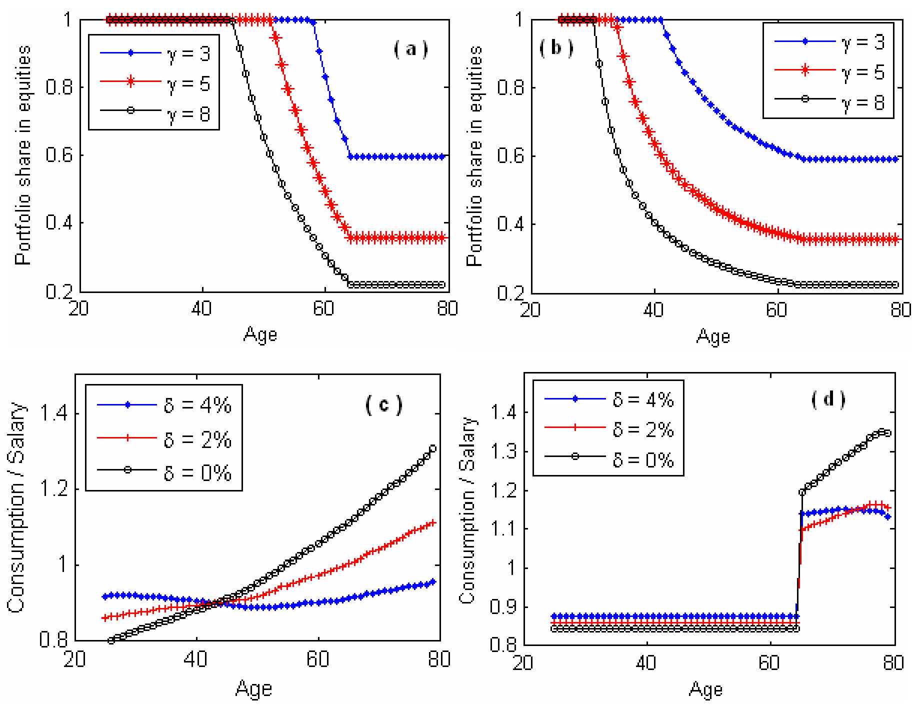
Since the investor is borrowing and short sales constrained, the optimal portfolio is then given by: $\omega_t = \min(\max(0, \omega_t^*), 1)$. The optimal contribution rate m is found by numerical search, given the optimal portfolio strategy.

2.8.3 Sensitivity analysis

The degree of risk aversion has a major impact on the portfolio choice. Figure B3 (upper panel) displays the average life cycle portfolio profiles for investors with different risk aversion $\gamma = 3, 5, 8$. Less risk averse investor allocates more assets in equities. The degree of risk aversion has only a minor impact on consumption and asset accumulation profiles, via the different portfolio choices. The more risk averse individuals save more and consume less compared with the less risk averse individuals. Figure B3 (lower panel) also shows the average life cycle consumption profiles for investors with different subjective discount rate $\delta = 4\%$, 2%, and 0%. The subjective discount rate has a strong impact on consumption profiles. In general, from the first order condition for optimality (Equation 2.33 in the appendix) it follows that the expected consumption is increasing when the expected portfolio return is larger than the subjective discount rate. The figures show that for $\delta = 4\%$, the median consumption profile is fairly flat. For lower values of δ the median consumption is increasing with age. The subjective discount rate has only a minor impact on the portfolio choice, via the changes in total-wealth to financial wealth ratio.

Figure B3: **The impacts of different risk aversion and subjective discount rates on individual schemes**.

Varying the degree of risk aversion among $\gamma = 3$, 5, and 8, while keeping other parameters at their default values, (a) shows the median of portfolio weight distribution in the OI scheme over ones life cycle, and (b) shows the median of portfolio weights distribution in the DC scheme. Varying the subjective discount rate among $\delta = 4\%$, 2%, and 0%, while keeping other parameters at their default values, (c) shows the median of consumption distribution in the OI scheme, and (d) shows the median of consumption distribution in the OI scheme.



2.9 Tables and Figures

Table 2.1: **Optimal pension schemes**
The table shows the optimal pension scheme parameters under the default values for the time preference ($\delta = 4\%$) and equity premium ($\mu - r = 4\%$). The optimal contribution rate and the corresponding replacement ratio are given by (p, b). The optimal speed of risk absorbing by adjusting contributions or benefits is determined by (α, β) respectively. The optimal portfolio weight in equities is given by (ω). The welfare levels achieved under these optimal collective schemes are indicated by CEC (in units of annual salary). The ratio CEC/CEC^{OI} shows the relative welfare gain or loss of the collective schemes relative to the optimal individual scheme (OI).

γ		OI	DC	DB_{CA}	DB_{BA}	DB_H
3	CEC	0.908	0.886	0.926	0.886	0.943
	CEC/CEC^{OI}	100%	97.6%	102.0%	97.6%	103.9%
	(p, b)	-	11.0%	16.6%,78.4%	11.4%,53.6%	14.0%,66%
	α	-	-	0.07	-	0.06
	β	-	-	-	0.03	0.02
	ω	-	-	100%	94%	100%
5	CEC	0.892	0.867	0.889	0.867	0.912
	CEC/CEC^{OI}	100%	97.2%	99.7%	97.2%	102.3%
	(p, b)	-	12.8%	16.6%,78.4%	13.1%,53.6%	14.0%,66%
	α	-	-	0.05	-	0.045
	β	-	-	-	0.03	0.02
	ω	-	-	96%	60%	100%
8	CEC	0.876	0.854	0.865	0.853	0.888
	CEC/CEC^{OI}	100%	97.7%	98.7%	97.3%	100.9%
	(p, b)	-	14.1%	17.5%,82.5%	14.0%,66%	15.7%,74.4%
	α	-	-	0.04	-	0.04
	β	-	-	-	0.02	0.02
	ω	-	-	67%	55%	76%

Table 2.2: **Welfare of the future generations**
This table reports the welfare improvements CEC^f/CEC^{OI} of the future generations who enter the optimal schemes (as in Table 2.1) in f years' time. Risk aversion equals $\gamma = 5$. The benchmark for this welfare comparison is set at the optimal individual scheme, OI. The future time of entering into the pension schemes is denoted by f.

	$f=0$	$f=2$	$f=5$	$f=10$	$f=30$	$f=100$	$f=1000$
DB_{CA}	99.7%	100.2%	101.7%	104.2%	112.3%	129.6%	149.5%
DB_{BA}	97.2%	97.3%	97.5%	97.8%	98.7%	100.7%	101.1%
DB_H	102.3%	102.8%	104.3%	106.6%	114.1%	126.7%	137.2%

Table 2.3: **Optimal pension schemes with lower equity premium**
The table shows the optimal pension scheme parameters for a reduced equity premium ($\mu - r = 3\%$). The optimal contribution rate and the corresponding replacement ratio are given by (p, b). The optimal speed of risk absorbing by adjusting contributions or benefits is determined by (α, β) respectively. The optimal portfolio weight in equities is given by (ω). The welfare levels achieved under these optimal collective schemes are indicated by CEC (in units of annual salary). The ratio CEC/CEC^{OI} shows the relative welfare gain or loss of the collective schemes to the optimal individual scheme (OI).

γ		OI	DC	DB_{CA}	DB_{BA}	DB_H
3	CEC	0.882	0.863	0.886	0.865	0.899
	CEC/CEC^{OI}	100%	97.7%	100.5%	97.8%	101.9%
	(p, b)	-	12.4%	16.1%,76%	12.2%,57.6%	14.3%,67.2%
	α	-	-	0.08	-	0.055
	β	-	-	-	0.03	0.025
	ω	-	-	92%	76%	95%
5	CEC	0.869	0.852	0.864	0.850	0.876
	CEC/CEC^{OI}	100%	97.9%	99.4%	97.8%	100.8%
	(p, b)	-	14%	17.5%,82.5%	14%,66%	14.3%,67.6%
	α	-	-	0.065	-	0.055
	β	-	-	-	0.025	0.02
	ω	-	-	62%	50%	88%
8	CEC	0.859	0.842	0.846	0.842	0.860
	CEC/CEC^{OI}	100%	98.0%	98.5%	98.0%	100.1%
	(p, b)	-	15%	18.9%,88.8%	14.8%,70%	15.7%,74%
	α	-	-	0.05	-	0.05
	β	-	-	-	0.02	0.02
	ω	-	-	45%	40%	62%

Table 2.4: **Market value of intergenerational transfers**
This table reports the market value of the actual contributions (PVP) and benefits (PVB) for a new entry cohort and an initially fully funded pension scheme ($FR_0 = 1$). The market value of the intergenerational transfers are shown in columns 'call' (for positive transfers) and 'put' (for negative transfers), as defined in equation (2.27). All scheme parameters are set at the scheme-specific optimal levels as in Table 2.1 for $\gamma = 5$. The market values are expressed in terms of multiples of annual salary.

$\gamma = 5$	p	PVP	PVB	call	put
DB_{CA}	16.6%	4.62	4.61	1.26	1.25
DB_{BA}	13.1%	3.64	3.64	0.22	0.22
DB_H	14.0%	3.89	3.88	0.77	0.77

Table 2.5: **Welfare evaluation of initially overfunded and underfunded schemes**
FR_0 denotes the initial funding ratio. The optimal scheme designs are as shown in Table 2.1 for $\gamma = 5$. CEC/CEC^{OI} shows the relative welfare gain or loss to the optimal individual scheme.

	OI	DC	DB_{CA}	DB_{BA}	DB_H
$FR_0 = 1.1$					
CEC	0.892	0.867	0.909	0.871	0.928
CEC/CEC^{OI}	100%	97.2%	101.9%	97.5%	104.0%
$FR_0 = 0.9$					
CEC	-	-	0.867	0.862	0.896
CEC/CEC^{OI}	-	-	97.4%	96.8%	100.7%
$FR_0 = 0.8$					
CEC	-	-	0.845	0.856	0.881
CEC/CEC^{OI}	-	-	94.8%	96%	98.8%

Table 2.6: **Market valuation of initially underfunded and overfunded schemes**
This table reports the market value of the actual contributions (PVP) and benefits (PVB) of the intergenerational transfers when the collective schemes are initially underfunded ($FR_0 = 0.8$, $FR_0 = 0.9$) or overfunded ($FR_0 = 1.1$). The schemes parameters are set according to Table 2.1, $\gamma = 5$. The market values are expressed in terms of multiples of annual salary.

$\gamma = 5$	p	PVP	PVB	call	put	NPV
$FR_0 = 1.1$						
DB_{CA}	16.6%	4.05	4.61	1.02	1.58	0.56
DB_{BA}	13.1%	3.64	3.79	0.15	0.30	0.15
DB_H	14.0%	3.55	3.91	0.62	0.98	0.36
$FR_0 = 0.9$						
DB_{CA}	16.6%	5.16	4.61	1.53	0.98	-0.55
DB_{BA}	13.1%	3.64	3.50	0.31	0.16	-0.15
DB_H	14.0%	4.21	3.86	0.94	0.60	-0.35
$FR_0 = 0.8$						
DB_{CA}	16.6%	5.71	4.61	1.75	0.65	-1.10
DB_{BA}	13.1%	3.64	3.35	0.42	0.12	-0.30
DB_H	14.0%	4.53	3.84	1.14	0.45	-0.70

Table 2.7: **Welfare of socially optimal collective schemes**
This table shows the social planner's optimal designs with weighting factor $\triangle = \delta = 0.96$ and risk aversion $\gamma = 5$. CEC^{opt} is the certainty equivalent consumption of the time 0 entry cohort under the optimal schemes chosen by the social planner.

	DB_{CA}	DB_{BA}	DB_H
CEC^{opt}/CEC^{OI}	98.7%	96.4	101.8%
(p, b)	19.2%,90.7%	14%,66%	15.7%,74.3%
α	0.04	-	0.03
β	-	0.025	0.01
ω	100%	65%	100%

Table 2.8: **Valuation of intergenerational transfers under socially optimal design**
This table reports the market value of the actual contributions, PVP, and benefits, PVB under the social planner's optimal designs with the weighting factor $\triangle = \delta = 0.96$ and risk aversion $\gamma = 5$. The market value of the intergenerational transfers are shown in columns 'call' (positive transfers) and 'put' (negative transfers).

$\gamma = 5$	p	PVP	PVB	call	put
DB_{CA}	19.2%	5.35	5.34	1.46	1.45
DB_{BA}	14%	3.87	3.87	0.3	0.3
DB_H	15.7%	4.36	4.35	0.98	0.97

Figure 1: **Life cycle profiles of the individual schemes (OI and DC)**

Taking a benchmark investor (with $\gamma = 5$, $\delta = 4\%$), the following charts show the life cycle quantile distributions of portfolio weights in equities for the *OI* scheme (a) and the *DC* scheme (b); the life cycle quantile distributions of consumptions for the *OI* scheme (c) and the *DC* scheme (d); and the life cycle quantile distributions of wealth accumulation for the *OI* scheme (e) and the *DC* scheme (f).

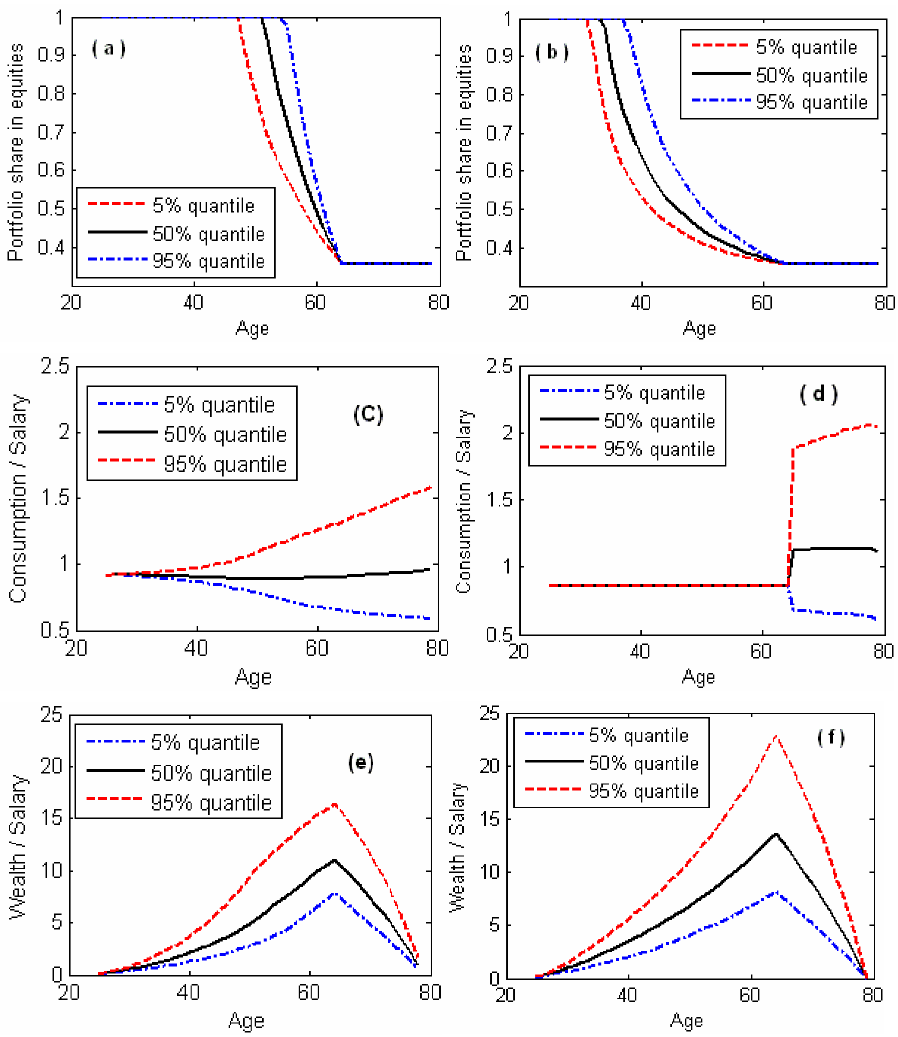


Figure 2: **Funding status of the collective schemes.**

The charts below show the (5%, 50%, 95%) quantile distribution of the surplus ratio, S_t/L, of the optimal collective schemes (DB_{CA}, DB_{BA}, and DB_H). The scheme parameters are fixed at their optimal values as given in Table 2.1 for $\gamma = 5$.

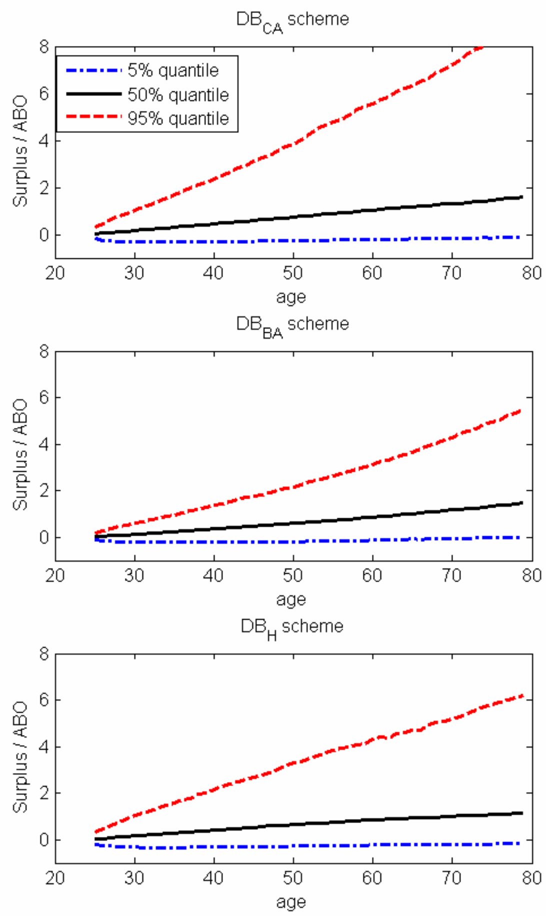


Figure 3: **Consumption profiles obtained from the collective schemes**

The charts below show the (5%, 50%, 95%) quantile distribution of consumptions in the optimal collective schemes (DB_{CA}, DB_{BA}, and DB_H). The scheme parameters are fixed at their optimal values as given in Table 2.1 for $\gamma = 5$. The consumption levels are normalized by annual salary, i.e. c_t/y.

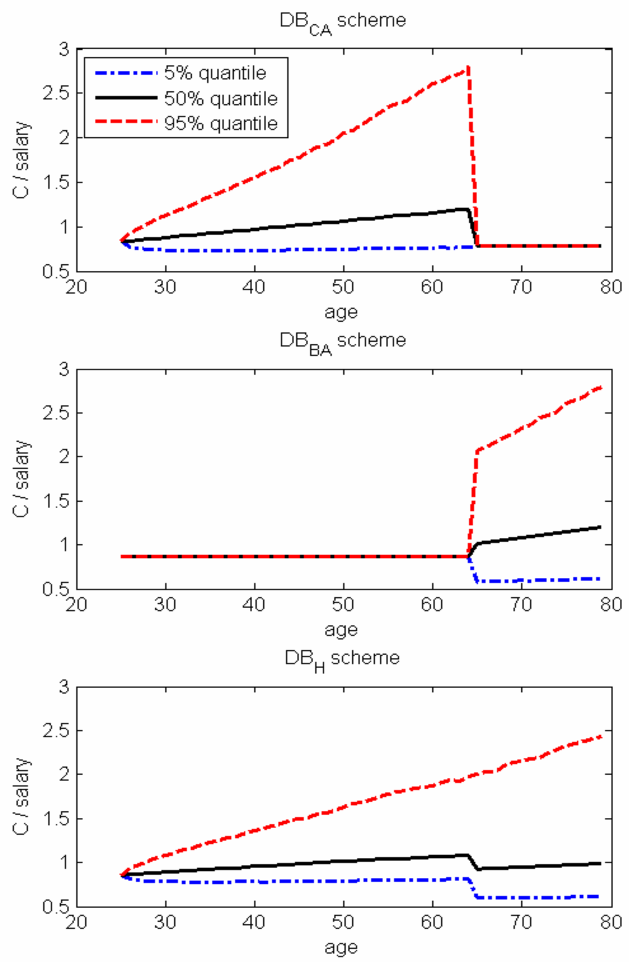


Figure 4: **Welfare levels of the living older generations**

This figure compares welfare levels of the living older generations under the optimal DB_H scheme (blue curve with dots, as in Table 2.1 for $\gamma = 5$), with what would be achieved under the OI scheme if the older generations switch to their own OI schemes (black curve).

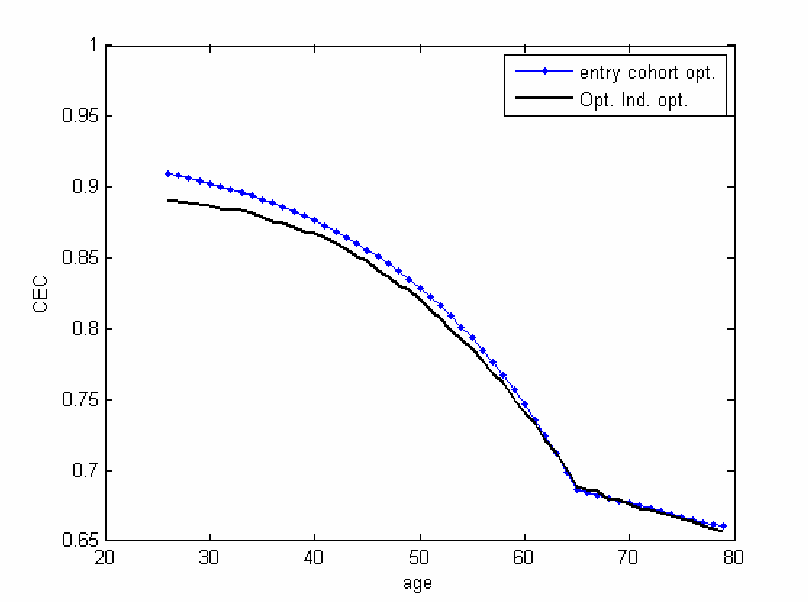


Figure 5: **Future generations under the generation-0 optimal schemes**.
Consider a future generation f who enters a given collective scheme in f years' time, and the scheme is optimized for the time-0 generation as characterized in Table 2.1 ($\gamma = 5$). The upper panel shows the welfare improvements obtained from entering the given collective schemes (DB_{CA}, DB_{BA}, and DB_H) above the optimal individual benchmark (OI), CEC^f/CEC^{OI}, for these 1000 future generations ($f = 1, 2, ..., 1000$). The lower panel shows the mean and the median of the surplus ratio (S_t/L) for the DB_H scheme.

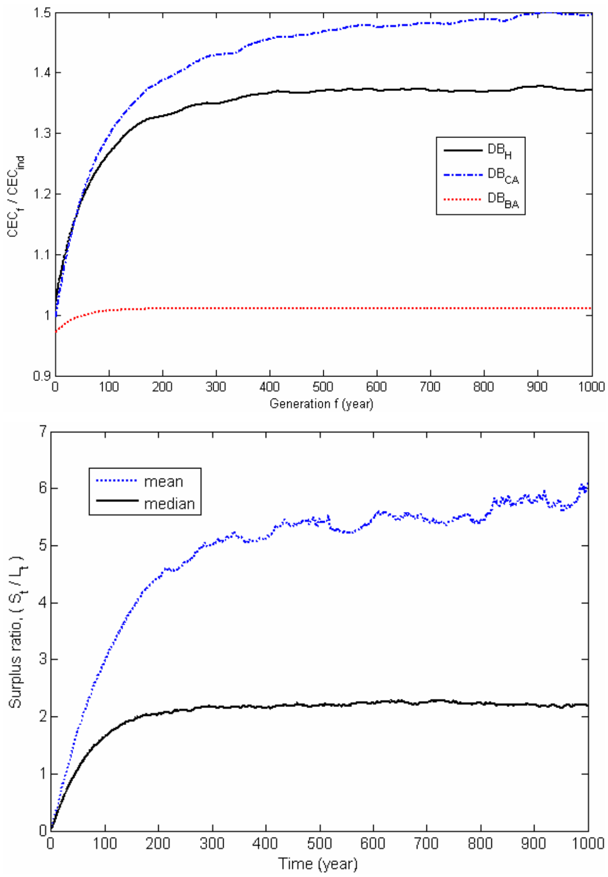
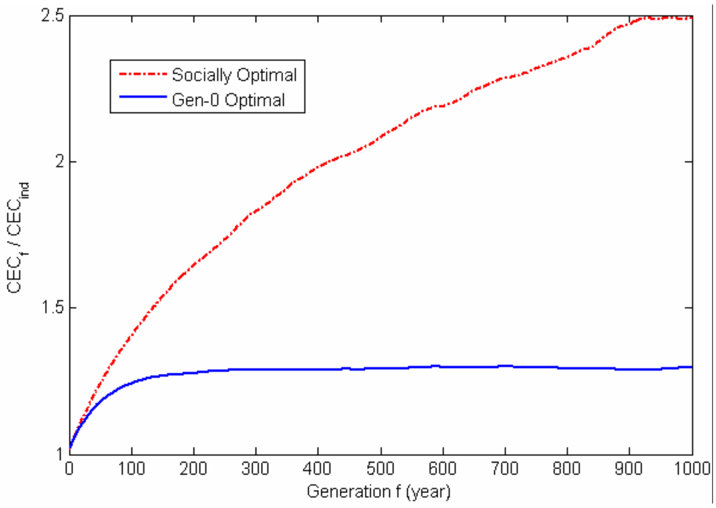


Figure 6: **Welfare gain under the socially optimal v.s. under the generation-0 optimal DB_H schemes**

This figure compares the welfare gain of the future generations between entering the socially optimal DB_H scheme (as in Table 2.7) v.s. the generation-0 optimal DB_H scheme (as in Table 2.1 ($\gamma = 5$)).



Chapter 3

DC pension plan defaults and individual welfare

This chapter is based on Cui (2008b).

3.1 Introduction

Individual DC pension schemes offer each participant the freedom to choose and to implement the optimal consumption and investment strategies to their own needs. Life cycle theory has shown how to determine the optimal saving and investing strategies. In reality, however, the early experiences with DC schemes show that most people choose highly suboptimal saving and investment strategies. A growing body of research shows that most people simply follow the given defaults[1] (Choi, Laibson, Madrian, Metrick (2004), Beshear, Choi, Laibson, Madrian (2004), Lusardi and Mitchell (2006), Benartzi, Peleg, and Thaler (2007)). They show that the default designs have significant impacts on the participation, contribution and investment outcomes. Choi, et al (2004) reported that, until late 1990s, non-participation was the standard enrollment default, i.e., the individuals are not enrolled unless one opts in. Under such default, participation rates were

[1] If individuals do not take action to choose from the avaiable options, then the default settings will be applied automatically. The defauls specify whether or not one contributes into the tax-deferred DC account, the level of contribution rate, and the investment funds in which the contributions will be invested.

low, ranging from 26-43% six month after hiring, and 57-69% three years after hiring. Since late 1990s, automatic enrollment started to be implemented in DC plans. Consequently, the reported participation rates exceeded 85% regardless of the tenure of the employee under automatic enrollment regime. Furthermore, 65%-87% of the participants adopted the default contribution rate of 3% or 4% of income, and the default investment in money market accounts. About 45% of the participants still stuck with these defaults three years later. Clearly, the default design has significant impact on retirement saving behavior.

Given the dramatic impact of defaults, naturally, we want the defaults to be as good as possible. Therefore, we ask the question: is the current popular default design the best possible design in welfare terms? If not, can we design a better default which may achieve a nearly optimal welfare outcome? To this end, the goal of this study is to find the optimal age-dependent contribution and investment rules, and to evaluate to what extent these default rules help to improve the individual welfare. We find, indeed, potentially large economic welfare gains by following the age-dependent defaults above the current standard default design. According to PSCA's *Automatic Enrollment 2001 Survey*[2] in the United States, the most common default contribution rate is 3% or 4% of pay (present in more than 60% of 401(k) plans). The most common investment default is stable value fund and money market fund (present in 67% of plans). In this paper, the current default design refers to automatic enrollment with flat contribution rate of 4% of income, together with a risk free investment vehicle.

The life cycle theory is a very useful framework for such analysis, and it provides us many insights[3]. There are two controls in the life cycle planning problems, namely the optimal consumption and portfolio choices. Both strategies may depend on one's age, income, wealth accumulation and other economic state vari-

[2] www.psca.org

[3] The life cycle theory has a long in finance literature. Particularly, Merton (1969, 1971), emphasizes on the role of human capital. Some of the more recent papers are Carroll (1992, 1994, 1997) on precautionary saving, Gourinchas and Parker (2002) on life cycle consumption, Viceira (2001), Gomes and Michaelides (2005) and Cocco, Gomes and Maenhout (2005) on life cycle portfolio choice, Benzoni, Collin-Dufresne, and Goldstein (2007) on life cycle strategies with cointegrated labor income with market returns, Cocco (2005) on portfolio choice in the presence of housing risk, and Gomes, Kotlikoff and Viceira (2008) on life cycle investing with flexible labor supply.

ables, and hence may differ by age and economic circumstances. A better default design thus should have two aspects, namely a better saving rate default and a better investment default. Several studies propose the so-called life-cycle funds as the portfolio allocation default (Bodie, McLeavey and Siegel (2007), Viceira (2007)). The idea of life-cycle funds is to mimic the theoretical life cycle portfolio strategies using an age-dependent portfolio rebalancing rule. In such life-cycle funds (also known as target maturity funds) the portfolio allocation to stock mutual funds declines as one ages, and is replaced gradually by safer assets like bonds and cash. In fact, the life cycle funds implement a simplified version of the optimal portfolio strategy. Life-cycle funds recently started to be implemented as the default by many DC scheme providers, and are expected to have large impact on the asset allocation outcome of most DC contributors in the future.

However, only changing the portfolio default alone does not help much if people do not contribute or fail to contribute enough. To address this pressing issue, recently, the United States passed the Pension Protection Act of 2006, which encourages the adoption of several 'autosave' features in the DC plans. These 'autosave' features include automatic enrollment, employer contribution, contribution escalation, and qualified investment default (see Beshears, Choi, Laibson, Madrian, and Weller (2008)). The contribution escalation is based on an interesting idea called Save More Tomorrow, which dramatically stimulates participants to save more (Thaler and Benartzi (2004)). In such a scheme, participants agree to automatically increase their saving rate whenever they receive a raise. However, as the authors claim, this design is solely based on behavior motivations, but not on financial or economic considerations. Are increased saving rates optimal or nearly optimal over life cycle? These questions are not addressed in the literature.

Life-cycle funds have tackled the asset allocation aspect of the life cycle theory but not the consumption aspect. Therefore, the next step forward is to extend the fixed contribution rate to include some age-dependent features. The main idea in this paper is to design simple default rules for DC contribution and investment to mimic the optimal consumption during one's life cycle. In addition, we evaluate to what extent these age dependent default contribution and investment rules are beneficial to the participants and can be recommended as default options for individual pension plans.

Our main findings are the following. First, we find large economic welfare gains by following the smart but simple age-dependent contribution rule above the current standard default design. Comparing to the current defaults, the age dependent defaults lead to 7.2% increase of the certainty equivalent consumption per year. Over 60 years (in adulthood), the welfare gain amounts to 2.78 times first year labor income. Using the fully optimal strategies as welfare benchmark, the current default design delivers maximally 92.7% of welfare relative to the optimal welfare level. Whereas, the age-dependent contribution and investment default design delivers more than 99% relative to the optimal welfare level. Therefore, the simple age-dependent contribution and investment rules can achieve nearly optimal welfare level.

Second, we find that the contribution (or saving) choice has larger impact on welfare than the portfolio choice does. As compared to the current defaults (with flat contribution rate of 4% and fully risk free investment), improving the contribution policy alone increases welfare from 92.7% to 97.1% of the optimal welfare level. However, improving the asset allocation alone only increases welfare from 92.7% to 95.2%. Here we show that setting the contribution (or saving) right is more important in welfare terms. The life cycle literature has been mainly focusing on the portfolio choices. However, we find the contribution rule plays a more important role in improving welfare.

Where does the welfare improvement come from? Our analyses reveal that it comes from a better trade-off between liquidity constraints and tax advantages. Early in life, individuals face liquidity constraints, because wage earnings are on average upward sloping over life time, but individuals cannot borrow against their future labor income to boost their consumption early in life. In addition, in data as well in our model setup, the housing expenditure is relatively high for the young than for the old, which make the liquidity constraint more binding for the young. On top of these, borrowing or early withdrawal from DC pension scheme is not allowed unless under severe circumstances and subject to high penalty costs (about 10% reduction). Therefore DC scheme is (nearly) illiquid during the whole working period[4], which makes the liquidity constraint more severe, especially early

[4]In the baseline model of this paper, we assume that borrowing or early withdrawal from the DC savings are not allowed.

in life. However, retirement saving through DC scheme is entitled to tax benefits (and often employer matches). As will be clear in the paper, tax benefits are higher as one ages. Thus, the age-dependent DC contribution rule avoids early periods when liquidity constraint is tight and tax benefit is low, but make best use of the later periods when liquidity is abundant and tax benefit is high.

Furthermore,we stress the idea of integrated retirement saving strategies, where retirement provision is considered jointly with important expenditures like housing and medicare. Therefore, we carefully model the dynamics and uncertainties of labor income, housing and medical expenditures, in order to realistically quantify and evaluate the default designs. As we will see that the decreasing life-cycle pattern of housing expenditures postpones DC savings in the beginning, but the increasing medical expenditures towards the end of life promotes retirement savings.

A closely related paper by Gomes, Michaelides and Polkovnichenko (2008) study the optimal portfolio choices and optimal saving strategies for rational individuals with a taxable account and a tax-deferred DC account.[5] Their focus is to match the calibrated life-cycle model with the empirical patterns in portfolio choice, wealth accumulation and stock market participation in the two accounts setting. This paper adapts a similar modeling setup as Gomes et al (2006), but with a focus on the optimal age-dependent contribution and investment rules. Gomes, Kotlikoff and Viceira (2008) also study the welfare comparisons of simple defaults. However they only consider defaults for portfolio choice.

The classical life cycle models with single account is inadequate for studying contribution rules. Therefore, the setting with two accounts is particularly important to capture the liquidity constraint when designing the contribution rules. In this paper, we explicitly model the liquid taxable account and illiquid tax-deferred DC plans. Under the two-account setting, we first illustrate the optimal contribution and investment policies. Then, we show various default designs, including constant and age-dependent features. We use the dynamic programming and the Endogenous-Grid Method (Carroll, 2006, 2007) to solve the extended life cycle

[5] Dammon, Spatt and Zhang (2004) also study the optimal portfolio choices with taxable and tax-deferred accounts. They focus on the portfolio choices in taxable and tax-deferred accounts due to different tax treatment on dividends, capital gains and interests.

model with two accounts.

The organization of the paper is the follows. Section 2 explains the model setup and the economic environment. Section 3 solves the life cycle model to describe the optimal life cycle saving and investment strategies in an ideal world. There we see several age related patterns regarding the asset allocation and consumption strategies over life time. In Section 4, we study the age-dependent default designs for passive participants in individual-based DC schemes. We focus on the effect of age-dependent designs of contribution and investment defaults, and to see how far one can push the welfare closer to the optimal strategies by using these age-dependent defaults. Section 5 presents the welfare comparisons and policy implications, and Section 6 concludes.

3.2 Model Setup

3.2.1 Individual's preferences

We assume that all individuals start working at age 25 ($t = 0$) and retire at 65 ($R = 40$). For simplicity, we assume that the individuals die at age 85 ($T = 60$). During the working period ($1 \leq t \leq R$) the individuals earn stochastic labor income, denoted by Y_t. During the retirement period ($R \leq t \leq T$), the individuals receive no income but consume their accumulated wealth, denoted by W_t. Individuals derive utility over single consumption goods (normalized by price inflation). An individual's preference is captured by the constant relative risk aversion utility function as

$$E\left[\int_{t=0}^{T} \beta^t \frac{C_t^{1-\gamma}}{1-\gamma} dt\right]$$

where γ is the risk aversion parameter and β is the subjective discount factor. In the baseline model, we fix $\gamma = 5$, and $\beta = 0.97$, following the life cycle literature.

3.2.2 Return dynamics

There are two financial assets traded in the market, one risk free and one risky (both in real terms). The real risk free asset offers a fixed real interest rate r. The real price of the risky stock index, E_t, follows geometric Brownian motion with a constant drift. Dividends are reinvested. The aggregate real wage index G_t is stochastic. The uncertain stock returns are potentially correlated with stochastic aggregate wage growth. This contemporaneous correlation is denoted by ρ, where $dZ_{E,t}$ and $dZ_{g,t}$ denote the independent Brownian incremental of stock returns and aggregate wage growth rates. The stock return and wage growth rate dynamics are the following

$$dE_t/E_t = \mu dt + \sigma_E\sqrt{1-\rho^2}dZ_{E,t} + \rho\sigma_E dZ_{g,t} \tag{3.1}$$

$$dG_t/G_t = \overline{g}dt + \sigma_g dZ_{g,t} \tag{3.2}$$

where μ and $\overline{g}$ are the instantaneous drifts, and σ_E and σ_g are the volatilities of stock returns and wage growth respectively. In the baseline model, we assume $\mu = 6\%$, $\sigma_E = 15\%$, $\overline{g} = 0.8\%$, and $\sigma_g = 4\%$ annualized.

3.2.3 Labor income and social security

Let t denotes calendar year and t_0 the year of birth, so that $t - t_0$ is the age of the individual under consideration. Following Benzoni, Collin-Dufresne, and Goldstein (2007), we assume that the individuals' real labor income $Y_{t-t_0} = G_t N_{t-t_0}$ can be decomposed into two component, an aggregate wage component, G_t, and an age-dependent idiosyncratic component N_{t-t_0}. The growth rate of aggregate wage component is determined according to eq(3.2).[6] While the idiosyncratic wage component, N, has an age-dependent drift $f(t - t_0)$ to generate the hump-shape of earnings and normally distributed permanent shocks. The real labor income process is specified as follows

[6] The aggregate wage growth and equity returns are cointegrated in Benzoni, Collin-Dufresne, and Goldstein (2007), hence the asset allocation to equities may be reduced due to the cointegration effect. Given the focus of this paper on the age-dependent contribution policy rule, we asset the aggregate wage rates and equity returns are correlated, but no cointegrated.

$$
\begin{aligned}
Y_{t-t_0} &= G_t N_{t-t_0} && (3.3)\\
\ln N_{t-t_0} &= \ln N_{t-1-t_0} + f(t-t_0) + \sigma_n \eta_t && (3.4)\\
&= \ln N_{t-1-t_0} + \left(a_0 + a_1\left(t-t_0\right)\right) + \sigma_n \eta_t && (3.5)
\end{aligned}
$$

where $\eta_t \sim i.i.d.N\left(0,1\right).$ The starting annual salary at age 25, N_{25}, is normalized to \$20,000. The parameter a_0 and a_1 are set according to the calibration of Benzoni, Collin-Dufresne, and Goldstein (2007) for the high education group ($a_0 =$ 0.066 and a_1= -0.0024), and $\sigma_n = 8\%$. Figure 1 shows the quantile distributions of N and G over time.

The old-age social security benefit at age 65, SS_R, is assumed to be a fraction, s, of the final labor income, i.e., $SS_R = s * Y_{R-1}$. Each year, the social security is indexed with the aggregate wage growth, so that $SS_t = G_t SS_R$, for $t > R$. In the baseline model, we consider $s = 30\%$.[7]

3.2.4 Housing and medical expenditures

From an integrated retirement saving point of view, retirement provisions should be considered in combination with important expenditures like housing and medicare, because these expenditures affect when and how much one should save for retirement provisions. As shown in Section 3, the decreasing life-cycle pattern of housing expenditures postpones DC savings in the beginning, but the increasing medical expenditures towards the end of life promotes retirement savings. In this paper, we carefully model the dynamics of housing and out-of-pocket medical expenditures based on the literatures. Housing expenditures exhibit a decreasing age profile (Gomes and Michaelides (2005), Amromin, Huang and Sialm (2007)) and medical expenditures exhibit an increasing age profile (Palumbo (1999), Scholz, Seshadri and Khitatrakun (2006)). Both expenditures are modeled as exogenous shocks to the budget process.

We assume that individuals pay off all their mortgages before age 80. The exogenous housing expenditure represents a fraction of labor income during working

[7] $s = 30\%$ might overstate the benefit for the high final salary individuals, and understate the benefit for the low final wage individuals.

period, and a fraction of final income during the retirement period. Based on the estimation of Gomes and Michaelides (2005) using PSID data, the ratio of housing expenditure to income has the following age-dependent mean and variances

$$H_t/Y_t = h_t \sim N\left(\overline{h}\left(t\right), \sigma_h^2\left(t\right)\right) \tag{3.6}$$

where the age-dependent mean $\overline{h}\left(age\right) = h_0 + h_1 * age + h_2 * age^2 + h_3 * age^3$, with $h_0 = 0.71$, $h_1 = -0.035$, $h_2 = 0.00072$, $h_3 = -0.0000049$. Furthermore, uncertainty in the housing expenditure is captured by $\sigma_h = 2\%$, and the correlation between shocks to h_t and income Y_t is set at -0.5. Figure 2 (left panel) depicts the average profile and one simulated scenario of the housing expenditures over life cycle. Housing expenditures are relatively high for the young and the mid-age households, but relatively low for the retired households when mortgages are gradually paid off.

The medical expenditure also represents a fraction of labor income during working period, and a fraction of final income during the retirement period. The ratio of out-of-pocket medical expenditure to income, following the parameterization and estimation results of Scholz, Seshadri and Khitatrakun (2006), is modeled as follows

$$\ln(M_t/Y_t) = -7.316 + 0.012 * age + 0.00066 * age^2 + \varepsilon_M \tag{3.7}$$

with $\varepsilon_M \sim N\left(0, \sigma_M^2\right)$ and $\sigma_M = 20\%$. Figure 2 (right panel) depicts the average profile and one simulated scenario of the medical expenditures over life cycle. Medical expenditures are more costly at advanced ages.

3.2.5 Tax-Deferred Account and Taxable Account

The tax-deferred account is actually an individual-based DC scheme that is provided by the employer in many countries. Tax advantages, and often additional employer matching, are used to stimulate savings in the DC schemes. These individual DC schemes are called the Tax-Deferred Account, since income tax and dividends tax are exempted or postponed till retirement. In this paper, the DC contributions are exempted from a higher income tax, $\tau^y = 30\%$; withdrawals from DC account are taxable at a lower income tax rate, $\tau^o = 20\%$, when retired. Sometimes, employers will match the contributions made by employees. The employer matching is a bonus if the employee contributes. The tax deferral and employer matching provide certain incentives for individuals to save in the tax deferred DC account. However, to avoid large tax arbitrage, the contribution rate is capped at 20%, so that maximally 20% of gross income can be contributed in DC plan in each year. The individuals thus may keep their retirement savings in these Tax-Deferred Account, and may also hold other private wealth in a Taxable Account.

3.3 Optimal life cycle strategies

This section studies the optimal life cycle planning problem of an individual with a tax-deferred account (TDA) and a taxable account (TA), under the realistic economic settings as in Section 2. Section 3.1. describes the life cycle model for an individual with tax-deferred DC account and taxable account. Section 3.2. shows the optimal strategies under the tax benefit driven setting (the baseline model). Section 3.3. presents the results with additional employer matching.

3.3.1 Optimization problem with TDA and TA

The optimization problem with a tax-deferred DC account and a taxable account is as follows: The Individual optimizes the life time utility by optimally choosing consumption, contributions into the DC scheme, and the asset allocations in tax-deferred DC account and taxable account. We focus on when and how much to contribute into the DC scheme. Let m_t denote the contribution rates into the DC pension plan. We assume that one can not withdraw the DC wealth before retirement, that is, $m_t \geq 0$ (the contribution rates must be non-negative). In practice, under severe circumstances early withdrawal from DC account is allowed (but subject to 10% penalty cost). In the baseline model of this paper, we assume that borrowing or early withdrawal from the DC savings are not allowed, therefore DC account is illiquid during the whole working period. The illiquid DC savings make the liquidity constraint potentially more severe.

Let W_t^τ denote an individual's wealth in the taxable account, and W_t^{DC} be the wealth in the tax-deferred DC account. $\widetilde{R}_{t+1}^e = E_{t+1}/E_t$ denotes the total return on equities and $R^f = \exp(r)$ denotes the real risk free rate. The fractions of assets invested in equities are denoted by α_t^τ (for the liquid saving) and α_t^{DC} (for the DC account) respectively. Given the focus of this paper on consumption and saving decisions, we make the simplifying assumption that dividends and capital gains are not taxed, and further assume the asset allocation in taxable and tax-deferred DC accounts are identical. The descritized optimization problem with TDA and TA can be formalized as following:

$$V = \max_{\{C_t, \alpha_t^\tau, \alpha_t^{DC}, m_t\}_{t=1}^T} E\left[\sum_{t=1}^T \beta^{t-1} \frac{C_t^{1-\gamma}}{1-\gamma}\right] \tag{3.8}$$

subject to, the wealth dynamics, before retirement (1$\leq t < R$), as

$$\begin{aligned} W_{t+1}^\tau &= (W_t^\tau - C_t - (1-\tau)m_t Y_t)\left(R^f + \alpha_t^\tau\left(\widetilde{R}_{t+1}^e - R^f\right)\right) && (3.9)\\ &\quad +(1-\tau^y)Y_{t+1} - H_{t+1} - M_{t+1} \\ W_{t+1}^{DC} &= \left(R^f + \alpha_t^{DC}\left(\widetilde{R}_{t+1}^e - R^f\right)\right)\ \left[W_t^{DC} + m_t Y_t\right] && (3.10)\\ \text{with } W_1^\tau &= (1-\tau)Y_1, \text{ and } W_1^{DC} = 0 && (3.11) \end{aligned}$$

Since the individuals are borrowing constrained, the balances of the two savings must always be non-negative, as in eq(3.12).

$$\begin{aligned} W_t^\tau &\geq 0,\ W_t^{DC} \geq 0 && (3.12)\\ 20\% &\geq m_t \geq 0 && (3.13) \end{aligned}$$

After retirement ($T \geq t \geq R$), DC savings become liquid and available for consumption. The individual thus combines the two savings after deducting the income tax paid on the DC wealth, i.e., $W_R = W_R^\tau + (1-\tau^o)\, W_R^{DC}$. These combined wealth are invested in the taxable account to finance the retirement consumption. Formally, the wealth dynamics of the savings during the retirement period are as follows.

$$W_{t+1} = \left[\alpha_t^{\tau} \widetilde{R}_{t+1}^{e} + (1-\alpha_t^{\tau}) R_t^f\right] [W_t - C_t] + (1-\tau^o) SS_{t+1} - H_{t+1} - M_{t+1} \quad (3.14)$$

Furthermore individuals are short-sales constrained, which implies that

$$1 \geq \alpha_t^{\tau} \geq 0 \ , \ 1 \geq \alpha_t^{DC} \geq 0 \quad (3.15)$$

The problem has no analytical solution due to the portfolio constraints. We use the dynamic programming principle together with the Endogenous-Grid Method (Carroll (2006, 2007)) to solve the extended life cycle model with two accounts. See Appendix A for the details of the solution technique.

3.3.2 The life cycle saving and investing profiles

This subsection shows the optimal life cycle profiles of the individuals with both taxable and tax-deferred DC accounts. The distribution of the life cycle profiles are characterized by 5%, 50% 95% quantiles.

Figure 3 (left panel) shows the portfolio allocation in stocks over life time. Recall that we assume the asset allocation in Taxable and Tax-deferred DC accounts are identical. As explained by Campbell and Viceira (2002), due to the leverage effect of human capital[8], the portfolio allocation to equities generally decreases over time. Here we confirm this finding under the two accounts setting. Figure 3 (right panel) shows the consumption profile, which is slightly increasing over time. It is due to the no-borrowing constraint and the assumed time preference.

Figure 4 (left panel) shows the contribution rate profile. Strikingly, the young individuals make zero contribution to the Taxable DC account for the first ten years of working period. The contribution rate rapidly increases since the mid-age, and eventually reaching to 20% ceiling a few years before retirement. Figure 4 (right panel) shows the wealth accumulation in the TA and TDA accounts. Early in life, young individuals accumulate moderate wealth in taxable account as precautionary savings against various background risks (incl. income uncertainties and expenditure uncertainties). The DC wealth starts to grow rapidly after mid-

[8] The human capital is the present value of the future incomes.

age, when the individuals are not liquidity constrained, their retirement saving motives get stronger, and tax benefits are much higher. The growth of TDA account is boosted by transferring wealth from TA to TDA account via higher contributions around mid-age.

3.3.3 With Employer Matching

Here we consider a variant of the baseline model, including the employer matching. As reported by Gomes, Michaelides and Polkovnichenko (2008), about half of the companies do not provide employer matching, and for the other half of the companies who do match employees' contributions, the employer matching may take various form in practice. Here we consider a common practice, i.e., the employer matches 100% of the employee's contribution up to a limit of 6%. However, the total contribution should not exceed 20%. Let m_t^{DC} denote the total contribution rates into the DC pension plan.

$$m_t^{DC} = \min\left(m_t + \min\left(m_t,\ 6\%\right),\ 20\%\right) \tag{3.16}$$

The DC wealth dynamic before retirement ($t < R$) is slightly modified to

$$W_{t+1}^{DC} = \left(R^f + \alpha_t^{DC}\left(\widetilde{R}_{t+1}^e - R^f\right)\right)\left[W_t^{DC} + m_t^{DC} Y_t\right] \tag{3.17}$$

Figure 5 to 6 show the life cycle profiles of portfolio choice, contribution rates, consumption and wealth accumulations, in the DC scheme with employer matching. Most of the profiles are very similar to those in the baseline setup. The main difference is in the contribution rates profile, Figure 6 (left panel). It exhibits a clear ladder shape, increasing over life cycle. The contribution rate is zero for the first 9 years of working life. The young individuals are liquidity constrained, because they face relatively low incomes but high housing expenditures. They leave the employer match on the table. Between age 35 and 45, the contribution rate increases quickly to 6%, earning the maximal employer matching. After age 45, it increases up to 14%, so that together with the employer matching the total contribution reaches the 20% ceiling.

3.3.4 Welfare evaluation

We use the certainty equivalent consumption (CEC) as the welfare measure, which can be easily backed out from the following equality

$$V = \sum_{t=1}^{T} \beta^{t-1} \frac{CEC^{1-\gamma}}{1-\gamma} \tag{3.18}$$

For easy interpretation, we divide the CEC by the first gross labor income Y_{25}. The welfare obtained by implementing the optimal strategies are the following. In our baseline model (i.e. no employer matching is provided), the optimal CEC reaches 0.688 (in units of first year income). In the variant where employer matching is offered, the CEC is higher, reaching 0.702 units of first year income. The welfare obtained in this section set upper limits for the comparison across various defaults in the next section.

3.4 Default designs

The previous section discusses the optimal strategies. This section discusses the default designs for the tax deferred DC accounts. Given the dramatic impact of defaults on retirement saving behavior, we want the default to be as good as possible. Is the current popular default design the best possible design in welfare terms? If not, can we design a better default which may achieve nearly optimal welfare outcome? For designing the optimal default options, we model the cases where individuals staying with the defaults throughout the life cycle, as if the default DC plan were made mandatory.

We consider four specifications of default designs. The main characteristics of these default designs are summarized in the following table.

	Default #1	Default #2	Default #3	Default #4
portfolio	constant	age-dependent	constant	age-dependent
contribution	constant	constant	age-dependent	age-dependent

Table 1: Overview of the default designs

Under each given default design specification, the individuals in the model optimize the their objective (3.8), subject to the budget constraints in TA and TDA (3.9, 3.10), and no-borrowing constraint (3.12) and no-short selling constraint (3.15). Detailed solution methodology is given in Appendix B. In Section 5, we will compare the welfare costs of different default designs relative to the optimal strategies, and investigating whether the age-dependent default is able to help to reduce the welfare cost.

3.4.1 Default #1: constant contribution rate and constant portfolio

This default is featured by a constant contribution rate and a constant portfolio choice throughout (working) life. These constant features resemble the current standard default options, which typically fix the contribution rate at 4% and invest in money market accounts, without further adjustment.

Individuals under this default design optimize utility over life time consumption, by choosing the consumption level in each period, and a constant saving rate m and a constant portfolio α at the beginning of their careers. Effectively, we are replacing the optimal strategies in Section 3 by constants, i.e., $\alpha_t^{DC} = \alpha_t^{\tau} = \alpha$, and $m_t^{DC} = m$ (with $0 \leq m \leq 20\%$). As before, borrowing and short sales are not allowed, which implies $0 \leq W_t^{\tau}$, and $0 \leq \alpha \leq 1$ respectively. The problem is solved numerically, by first optimizing the welfare level for given levels of contribution rate, and then determining the best contribution rate. Details of the solution procedure is given in Appendix B.

Table 2 shows, for a given flat contribution rate, the corresponding optimal fixed portfolio choice and the welfare level under the specification of Default #1. It shows that if contribution rate is fixed at 4%, then the best portfolio is investing 82.5% in equities throughout life cycle. This gives an annual certainty equivalent consumption of 0.654 units of first year labor income; Or, relative to the fully optimal TDA / TA benchmark (i.e., the upper limit) $CEC^{TDATA} = 0.6885$, it gives 95% of the optimal welfare level. Table 2 also shows the results for two additional asset allocations: one with 50% in equities and 50% in risk free asset, another case with 100% in risk free asset (as in the current defaults). The 50%/50% asset mix reduces the welfare levels slightly, for each given contribution rate. But the 100% risk free investment strategy is clearly sub-optimal, resulting in a large welfare loss. The current default with a fixed contribution rate at 4% and invest in money market accounts gives maximally 92.7% of the optimal welfare benchmark level.

When comparing the CEC across different contribution rates, we find that, for the assumed amount of tax benefit, without employer matching, the zero contribution is the best outcome. It means that the tax benefit in our baseline model is not large enough to compensate for the liquidity loss when young. Because the individual can not borrow from the future labor income, the contribution rates early in life is too high, so that the liquidity constraint become more severe.

Table 3 shows the variant with employer matching. In this case, the employee contributes m per cent of income, and the employer matches the contributions with a cap at 6% as in (3.16). Table 3 shows that for a contribution rate of 4%, the optimal asset mix is 82.5% in equities. This gives an annual certainty equivalent consumption of 0.664 units of first year labor income; relative to the fully optimal TDA / TA benchmark $CEC^{TDATA} = 0.702$, it gives 94.6% of the optimal welfare level. When comparing the CEC across different contribution rates, it seems that, the employer matching together with the tax deferrals stimulate DC savings. Our model suggests that a small but positive contribution rate of 1% is beneficial for the individual.

3.4.2 Default #2: constant contribution rate and age-dependent portfolio

Changing from Default #1, we relax the restriction of a constant portfolio throughout life, but replacing it with an age-dependent strategy. One popular rule describes a linearly relationship between age (denoted by $t - t_0$) and the share of risky assets as

$$\omega_t = \max[0, \quad \min[f_0 + f_1 (t - t_0), 1]] \tag{3.19}$$

The individual follows the the age-dependent allocation rule. Individual optimizes his life time utility of consumption by choosing consumption in each period, and choosing the constant saving rate m and the parameter (f_0, f_1) at the beginning of his career. As before, borrowing or short sales are not allowed, which implies

$0 \leq \omega_t \leq 1$ and $0 \leq m \leq 20\%$.

Figure 7 shows the optimized age-dependent portfolio rule corresponding to a given contribution rate of 4%. This life cycle fund resembles the age profile of the optimal portfolio choices shown in Figure 3. Figure 8 shows the resulting life cycle quantile profiles of consumption (left panel) and the wealth accumulation is both TA and TDA accounts (right panel). Comparing to the optimal strategies in Section 3, the consumption in this default is slightly lowered for the young, and slightly higher for the retirees, because of the constant contribution rate of 4%. Since the contribution rate is fixed at 4%, the accumulated DC wealth at the end of the working period is substantially smaller than the one in the optimal situation.

3.4.3 Default #3: age-dependent contribution rate and constant portfolio

Changing from Default #1, we relax the restriction of a constant contribution rate throughout life, but still keep the constant portfolio in both TA and TDA. Therefore the main features of the TDA are the age-dependent contribution rate and a constant portfolio choice. A simple age-dependent contribution rate may depend linearly on age as follows:[9]

$$m_t = \max[0, \quad \min[d_0 + d_1(t - t_0), \quad 20\%]] \tag{3.21}$$

Suppose the individual follows the age-dependent contribution rule. The individual optimizes his life time utility of consumption by choosing consumption in each period, and choosing the constant portfolio α and the parameter (d_0, d_1) for the contribution schedule at the beginning of his career. As before, borrowing and short sales are not allowed, which implies $0 \leq \alpha \leq 1$ and $0 \leq m_t \leq 20\%$.

Figure 9 shows the optimized age-dependent contribution rule and portfo-

[9] An alternative modeling of the age-dependent contribution rate may include the quadratic term in age. For example

$$m_t = d_0 + d_1 * (t - t_0) + d_2 * (t - t_0)^2 \tag{3.20}$$

This alternative modeling is able to capture the possible hump shape of the optimal contribution rates as seen in Section 3. This specification will be investigated in the future research.

lio choice (which is 85%) for Default #3. The age-dependent contribution rule resembles the optimal life-cycle contribution profile as in Figure 3 (Section 3). Following this default, the individuals do not contribute to the DC plan during the first 12 years of working period, because of the liquidity constraints. Between age 38 and 60, the contribution rates are non-zero and increase linearly with age, by about 0.8% per year. Between age 60 and 64, the individuals are not liquidity constrained any longer, therefore the maximum contributions are optimally chosen, driven by the tax benefits. Figure 10 shows the resulting life cycle quantile profiles of consumption (left panel) and the wealth accumulation is both TA and TDA accounts (right panel). We see that DC wealth is accumulated rapidly after mid-age.

3.4.4 Default #4: age-dependent contribution rate and age-dependent portfolio

Default #4 is a combination of Default #2 and #3. It specifies an age-dependent contribution rule and an age-dependent portfolio rule, i.e.,

$$m_t = \max[0, \quad \min[d_0 + d_1(t - t_0), \quad 20\%]] \tag{3.22}$$

$$\omega_t = \max[0, \quad \min[f_0 + f_1 (t - t_0), 1]] \tag{3.23}$$

Figure 11 shows the optimized age-dependent contribution rate (left panel) and age-dependent portfolio rule (right panel) of Default #4 (no employer matching). There is no employer matching provided. The age-dependent contribution rate profile is very close to that obtained under Default #3 in Figure 9. The individual starts contributing into the tax-deferred DC scheme after age 35, and increases the contribution rate each year till age 55, when the contribution cap is reached. The age-dependent portfolio profile shows a similar decreasing pattern, where equity exposure starts to decline from 100% around age 50, till 60% at the end of life. The resulting consumption and wealth accumulation profiles are very close to the ones shown in Figure 4 as in the optimal benchmark setup.

Figure 12 (left panel) shows the optimized age-dependent contribution rate of Default #4 when employer matching is provided. The solid line depicts the contribution rate made by the employee and the dashed line depicts the total contribution rate together with the matching. The individual starts contributing into the tax-deferred DC scheme after age 35, and increases the contribution rate each year till age 50, when the total contribution rate reaches the ceiling of 20%. The age-dependent portfolio profile shows a similar decreasing pattern but with slightly higher exposure to equities. The equity exposure starts to decline from 100% around age 55, till 70% at the end of life.

3.5 Welfare comparisons

3.5.1 Baseline model

Table 4 summarizes the main findings of this paper. Table 4 compares the optimal strategies with the current default design, and the step by step improvements above it. The first row ("TDA & TA") reports the certainty equivalent consumption (CEC) of the optimal strategies, which sets an upper limit of the welfare level. The welfare measure, CEC, is expressed in units of first-year gross income. The second row ("Current Default") reports the CEC obtained under the current default setting of 4% contribution rate and zero equity exposure ($\alpha = 0$). The current default maximally reaches 92.7% of the optimal welfare benchmark.

Starting from the current situation, we first improve the asset allocation to a portfolio with 50% in equities, while keeping the contribution rate at 4%. This step improves the welfare from 92.7% to 94.8% of the optimal level, as in the third row. When further improve the asset allocation to an optimal level of 82.5% in equities, keeping the contribution rate at 4%, the welfare is slightly improved to 95% of the optimal benchmark, as reported in the forth row. Then, we further improve the asset allocation by choosing an optimal life cycle fund, using the specification of Default #2. This step further improves the welfare to 95.2%, as reported in the fifth row. We see that, the completely risk free asset allocation is clearly suboptimal for a reasonable risk averse individual ($\gamma = 5$). Having a naive portfolio (e.g., 50% / 50%) improves the welfare by 2%. Having an optimal age-dependent portfolio only further enhance the welfare slightly (by another 0.2%).

Now, instead of improving the asset allocation, we replace the flat contribution rate by an optimal age-dependent contribution rate (as specified in Default #3), with a portfolio investing 85% in equities. This step further improves the welfare by 4%, up to 99.2% of the optimal welfare level, as reported in the seventh row. Compared to the current default, the total welfare gain amounts to 2.7 times of first year income over an adulthood of 60 years.

In the last row, when having both contribution and portfolio defaults age-dependent (as in Default #4), the welfare comes very close to the optimal level, i.e., reaching 99.4% of the optimal level. Similarly, the welfare improvement obtained from the age-dependent portfolio rule above the optimal fixed mix is small (about 0.2%).

Another way to improve from the current defaults is by first changing the contribution rates to age-dependent rule, while keeping the asset allocation in risk free asset, i.e., $\alpha = 0$. The CEC is immediately improved to 0.668 from 0.638, as given in the sixth row. Improving the contribution policy alone increases welfare from 92.7% to 97.1% of the optimal level, while improving the asset allocation alone increases welfare from 92.7% to 95.2%. We see that contribution policy plays a bigger role in improving the welfare.

Figure 13 shows additional results for the sixth and seventh row (Default #3) as in Table 4. The figure depicts the two age-dependent contribution rate rules which correspond to two different asset allocations, i.e., 100% risk free investment v.s. the optimal asset mix of $\alpha = 85\%$, based on the baseline model assumptions ($\gamma = 5$, $\beta = 0.97$). The two contribution defaults are close to each other, meaning that the optimal age-dependent contribution rule is robust with respect to portfolio mix. The difference in contribution rates is small, maximally 2%. Figure 13 shows that slightly higher contribution rates are necessary to compensate for the low equity exposure (the lower expected returns). The resulting welfare loss is 2% as reported in Table 4 (the sixth and seventh rows).

The variants of DC schemes with employer matching give very similar welfare results, as reported in Table 5. The findings observed from the baseline model still hold in the setting with employer matching.

Our main findings are the following. First, the simple age-dependent contribution rule and appropriate investment strategies can achieve nearly optimal welfare level. We find potentially large economic welfare gains by following the simple age-dependent contribution rule above the current standard default design. Using the fully optimal strategies as welfare benchmark, the current default design delivers maximally 92.7% of welfare relative to the optimal welfare level. Whereas, the age-dependent contribution and investment default design delivers up to 99.4% of the optimal welfare level. Comparing to the current defaults, the age dependent defaults lead to 7.2% increase of the certainty equivalent consumption per year (which is 0.046 units of first year labor income). Over 60 years (in adulthood), the welfare gain amounts to 2.78 times first year labor income.

Second, we find that the contribution (or saving) choice has larger impact on welfare than the portfolio choice does. Comparing to the current default (with $\alpha = 0, m = 4\%$), improving the contribution policy alone (keep $\alpha = 0$) increases welfare from 92.7% to 97.1% of the optimal level; while improving the asset allocation alone (with $m = 4\%$) increases welfare from 92.7% to 95.2%. Another example shows a similar result: Replacing the optimal fixed asset mix by an optimal life cycle fund, the welfare is only increased only 0.2%. The life cycle literature has mainly focused on the importance of portfolio choices. Here we show that setting the contribution (or saving) right is more important in welfare terms.

3.5.2 Sensitivity analyses

We want to know how sensitive the age-dependent contribution and investment rules are with respect to different model assumptions, e.g. risk and time preferences, income profile, life expectancy, etc. If the age-dependent default rules are not very sensitive to the assumptions, then the defaults are applicable for heterogeneous participants. Otherwise, we should better tailor make the defaults by first characterizing the participants by questionnaires.

Risk and time preference

What are the optimal strategies for a $\gamma = 8$ individual? Figure 14 and 15 show the age profiles of the optimal contribution rates, optimal portfolio choices, and the resulting consumption and wealth accumulation. The optimal welfare for the

$\gamma = 8$ investor is $\text{CEC}^{TDATA,\gamma=8} = 0.582$. The $\gamma = 8$ individual starts contribution after age 30 (which is 5 years earlier than a $\gamma = 5$ individual does as in Figure 3), and then gradually increases the contribution rate till the maximal level. The portfolio choice is also more conservative since mid-age, where equity exposure is 20-30% lower as compared to $\gamma = 5$.

Figure 16 compares the optimal age-dependent contribution rate for a more risk averse individual ($\gamma = 8$) with the benchmark $\gamma = 5$ individual. The age-dependent contribution rule in Figure 16 (according to Default #3) for the $\gamma = 8$ individual approximates the age profile of the optimal contribution rates in Figure 14. The corresponding asset mix (according to Default #3) for the $\gamma = 8$ individual consists 55% in equities, compared to 85% of equities investments preferred by the $\gamma = 5$ individual. We fix other parameters the same as the baseline model. We observe that the contribution rule for the more risk averse individual is 4% higher than that for the $\gamma = 5$ individual. Also, the optimal strategy suggests that the more risk averse individual should start contributing 5 years earlier than the $\gamma = 5$ individual does.

What happens if a $\gamma = 8$ individual steps into a scheme where the age-dependent contribution rule is designed for $\gamma = 5$ individuals, but the asset allocation is fixed at the right level (i.e., 55% in equities)? The resulted CEC in this case is 0.571, which is 98.1% of the optimal level. The welfare loss is about 2% compared to the optimal welfare level. The cumulative welfare loss amounts to 0.6 unit of first year salary over a life time.

As to the sensitivity to the time preference parameter β, we find that both contribution and asset allocation rules are not very sensitive to different values for β, e.g. $\beta = 0.95, 0.97, 0.99$. To save space, the results are not shown here.

To conclude, we find that the age-dependent contribution rate should be higher and asset allocation should be more conservative for a more risk averse individual. The welfare cost for a more risk averse individual to follow a contribution rule that is designed for a less risk averse individual might be sizable. Therefore, it is better to first use questionnaire to characterize their risk preferences and design the age-dependent defaults for each group.

Wage earning profile

The baseline model assumes an individual with a steeper upward sloping earning profile, as the case for higher educated individuals for example. The average earning profile for lower educated individuals is flatter. How does the flatter earning profile affect the default designs? We set parameter a_0 and a_1 according to the calibration of Cocco, Gomes and Maenhout (2005) for the low education group, while keeping other parameters as the baseline setup. Figure 17 and 18 show the optimal strategies for individuals with flat earning profile. We see that the portfolio strategy is very close to that of a steeper earning individual (in Figure 3). However the contribution strategy starts about 6 years earlier and increases more gradually than the steeper earning case. Because the life time income is smaller for the lower educated group, the consumption and wealth quantile profiles are lower.

The age-dependent contribution and portfolio default rules mimic the optimal strategies closely (figures not shown here to save space). We notice that, the age-dependent contribution rule for flat wage earners (e.g. lower educated individuals) can start a few years earlier and gradually increase over time.

Life expectancy

The baseline model has assumed a life span of 85 years, with 40 years of working and 20 years of retirement. What if the life span is increased to 95 years, with 40 years of working and 30 years of retirement? Figure 19 (left panel) shows the optimal contribution strategy for the case with prolonged retirement period. The optimal DC contribution starts around age 35, which is the same as the baseline result, to avoid the liquidity constrained period. However, the contribution rates increase much rapidly afterwards. This is because, the individual has to save more to finance the longer retirement period. Figure 19 (right panel) shows the resulting wealth accumulations in TA and TDA. The optimal portfolio strategy and consumption profiles are very similar to the baseline results, hence the figures are not shown here. The age-dependent contribution default is similar as to the baseline result in Figure 11 (left panel), but with a steeper slope (figure not shown to save space). Therefore, when designing the age-dependent defaults, it is important to take the updated (population) life expectancy into account. The current uniform default has no concern about life expectancy at all, which is very inappropriate.

3.6 Conclusions

Given the dramatic impact of defaults in individual DC schemes, we investigate whether or not and how much we can improve the welfare by changing the default design. We find potentially large economic welfare gains by following the simple age-dependent contribution rule above the current standard default design. Using the fully optimal strategies as welfare benchmark, the current default design delivers only 92.7% of welfare relative to the optimal welfare level. Whereas, the age-dependent contribution and investment default design delivers up to 99.4% of the optimal welfare level. In terms of certainty equivalent consumption, the age dependent default leads to 7.2% increase in annual consumption. Therefore the simple age-dependent contribution and investment rules can achieve nearly optimal welfare level.

Second, we find that the contribution (or saving) choice has larger impact on welfare than the portfolio choice does. As compared to the current defaults (with flat contribution rate of 4% and fully risk free investment), improving the contribution policy alone increases welfare from 92.7% to 97.1% of the optimal welfare level. However, improving the asset allocation alone only increases welfare from 92.7% to 95.2%. Here we show that setting the contribution (or saving) right is more important in welfare terms. The life cycle literature has been mainly focusing on the portfolio choices. However we find the contribution rule plays a more important role in improving welfare.

As to the sensitivity of the defaults and the related implementation issues, we show that, it is better to use questionnaire to categorize the risk preferences and design the age-dependent defaults for each group. Furthermore, it is important to take the updated (population) life expectancy into account when designing the age-dependent contribution defaults.

As to the policy implications, the age-dependent contribution and investment rules can be recommended as default. Our finding is consistent with the auto-save features encouraged by the Pension Protection Act of 2006. This paper contributes to the life cycle literature and the DC industry by characterizing the optimal age-dependent contribution and investment rules for DC participants. We show that contribution policy plays a bigger role in improving the welfare.

3.7 Appendix A: Solution method of Section 3, with TA and TDA

We use the dynamic programming and the Endogenous-Grid Method (Carroll (2006, 2007)) to solve the extended life cycle model with two accounts. In the appendix, the consumption (C_t), wealth ($W_t^{\tau}, W_t^{DC}, W_t$) and expenditures ($H_t, M_t$) are all normalized by income Y_t. The normalized variables are denoted in small letters throughout. The state variables in this model are time and two wealth processes.

3.7.1 Solving the retirement period

First, we solve the retirement periods ($R \leq t \leq T$). Upon retirement, the taxable savings and tax-deferred DC savings are combined into one savings, because both are freely accessible for consumption. The combined wealth is denoted by $W_R = W_R^{\tau} + (1-\tau^o) W_R^{DC}$. A lower tax rate τ^o is applied to the DC wealth at the retirement date. Formally, the normalized objective function and the normalized budget constraint, for the retirement period $R \leq t \leq T$, are

$$
\begin{aligned}
v(w_t) &= \max \frac{c_t^{1-\gamma}}{1-\gamma} + \beta E_t \left[\left(R_{t+1}^G R_{t+1}^N \right)^{1-\gamma} v(w_{t+1}) \right] && (3.24) \\
s.t.\ w_{t+1} &= (w_t - c_t) \left[R^f + \alpha_t \left(\widetilde{R}_{t+1}^e - R^f \right) \right] \left(R_{t+1}^G \right)^{-1} && (3.25) \\
& \quad + (1-\tau^o) ss - h_{t+1} - med_{t+1} && (3.26) \\
w_t^{\tau} &\geq c_t; \quad 0 \leq \alpha_t \leq 1 && (3.27)
\end{aligned}
$$

The optimal consumption and portfolio choice $\omega_t^*(a)$, $c_t^*(a)$ and the endogenous optimal wealth $w_t^*(a)$ (for $a = \{a_j\}_{j=1}^J$) can be found by following procedure. The procedure starts by defining a new variable, $a_t = w_t - c_t$, as the after-consumption-wealth. Construct a grid of $a_t = \{a_j\}_{j=1}^J$. Now we solve for the optimal consumption and portfolio policies for each given a_j. The first order conditions w.r.t. α_t and c_t are

$$
\begin{aligned}
0 &= \beta E_t \left[v'(w_{t+1}) \left(\widetilde{R}^e_{t+1} - R^f \right) \left(\widetilde{R}^G_{t+1} \right)^{-\gamma} \right] && (3.28) \\
u'(c_t) &= \beta E_t \left[v'(w_{t+1}) \left(R^f + \alpha_t \left(\widetilde{R}^e_{t+1} - R^f \right) \right) \left(\widetilde{R}^G_{t+1} \right)^{-\gamma} \right] && (3.29)
\end{aligned}
$$

The Envelope theorem implies that $u'(c_t) = v'(w_t)$, since

$$
v'(w_t) = \beta E_t \left[v'(w_{t+1}) \left(R^f + \alpha_t \left(\widetilde{R}^e_{t+1} - R^f \right) \right) \left(\widetilde{R}^G_{t+1} \right)^{-\gamma} \right] \tag{3.30}
$$

Replace $v'(w_{t+1})$ by $u'(c_{t+1})$ in the two first order conditions, we have

$$
\begin{aligned}
0 &= \beta E_t \left[u'\left(c^*_{t+1}[w_{t+1}] \right) \left(\widetilde{R}^e_{t+1} - R^f \right) \left(\widetilde{R}^G_{t+1} \right)^{-\gamma} \right] && (3.31) \\
& c^*_t(a_t) && (3.32) \\
&= I_u \left(\beta E_t \left[u'\left(c^*_{t+1}[w_{t+1}] \right) \left(R^f + \alpha^*_t(a_t) \left(\widetilde{R}^e_{t+1} - R^f \right) \right) \left(\widetilde{R}^G_{t+1} \right)^{-\gamma} \right] \right)
\end{aligned}
$$

where $c^*_{t+1}[w_{t+1}]$ as the optimal consumption policy at time $t+1$, and $I_u(\cdot)$ denotes the inverse function of $u'(c_t)$. Using any numerical solver, Eq (3.31) will give the optimal portfolio weight $\alpha^*_t(a_t)$ for any given amount of investment a_t. Because of the borrowing and short-selling constraints, we then impose the restriction $0 \leq \alpha^*_t \leq 1$. Then, eq (3.32) will give the corresponding consumption $c^*_t(a_t)$ for any given amount of investment a_t. Finally, the optimal wealth process is endogenously determined by $w^*_t = c^*_t(a_t) + a_t$. The advantage of this method is that the numerical search is only needed once in solving $\alpha^*_t(a_t)$, while $c^*_t(a_t)$ can be directly obtained from (3.32).

Hence we obtain the corresponding policy function $c^*_t(w_t)$ for $t \geq R$. The value obtained at time R can be decompose into two terms $v(w_R) = u(w_R) K(T-R)$ where $K(t) = \frac{1}{F}(1 - \exp(-Ft))$, and $F = \frac{\delta - r(1-\gamma)}{\gamma} - \frac{(1-\gamma)(\mu - r)^2}{\gamma^2 \sigma^2}$. [10]

[10] After retirement, since there is no labor income nor social security available for the individual, the model is the classical Merton (1969) model. Without any portfolio constraint, the value function time time R has the following expression $v(w_R) = \frac{w^{1-\gamma}}{1-\gamma} K(T-R)$, as shown in Merton

Then, we solve the working periods ($1 \leq t < R$). But before moving backward into the working period, we need to map the vector of the single state variable $\{w_R^*(j)\}_{j=1}^J$ into two state variables with $w_{i,j}^\tau = w_R(j)\frac{i}{I}$, $w_{i,j}^{DC} = w_R(j)\frac{I-i}{I}(1-\tau^o)^{-1}$, with $i = 0, 1, ..., I$. This step is because both taxable and DC savings are state variables for the working period optimization problem. In a similar way, we map the vector $\{c_R^*(j)\}_{j=1}^J$ into a matrix with $c_{ij}^* = c_R^*(j)$ for $\forall i = 0, 1, ..., I$. Hence we obtain the corresponding policy function $c_{R,i,j}^*\left(w_{R,i,j}^\tau, w_{R,i,j}^{DC}\right)$ at time $t = R$.

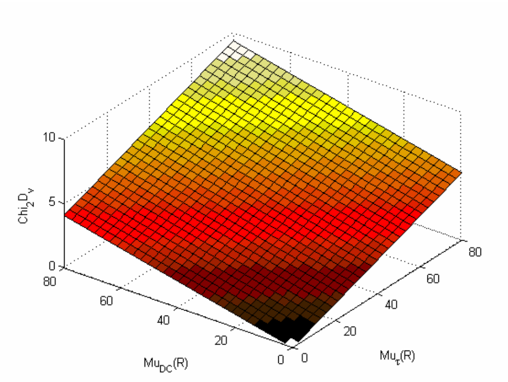


Figure: optimal consumption policy
$c_R^*\left(w_R^\tau, w_R^{DC}\right)$ at time $t = R$.

3.7.2 At time $t = R - 1$

At time $t = R-1$, the individual has two accounts, TA and TDA. Therefore the normalized value function has two state variables, $w_{R-1}^\tau, w_{R-1}^{DC}$. The individual has to decide how much to consume, c_{R-1}, out of the TA wealth, and where to locate his savings among the two accounts (by choosing m_{R-1}^{DC}); and finally the individual has to decide the right portfolio's $(\alpha^\tau, \alpha^{DC})$ for both TA and TDA.

Formally, the normalized value function and the normalized budget constraint are

(1969) and Munk (2007). With portfolio constraint (e.g. no-borrowing constraint), Grossman and Vila (1992, proposition 3.2.) show that the value function has the expression $v(w_R) = \frac{w^{1-\gamma}}{1-\gamma}\overline{K}(T-R)$, with $\overline{K}(t) = \exp\left(r + \mu k - A\sigma^2 k^2/2\right)(1-A)t$.

$$\varpi\left(w^{\tau}_{R-1}, w^{DC}_{R-1}\right) = \max_{c_{R-1},\alpha^{\tau},\alpha^{DC},m^{DC}_{R-1}} \frac{c^{1-\gamma}_{R-1}}{1-\gamma} + \beta E_{R-1}\left[\left(R^G_R R^N_R\right)^{1-\gamma} v\left(w_R\right)\right] \quad (3.33)$$

s.t. the budget constraint

$$\begin{aligned} w_R &= w^{\tau}_R + (1-\tau^o)\, w^{DC}_R && (3.34)\\ w^{\tau}_R &= \left(w^{\tau}_{R-1} - c_{R-1} - (1-\tau^y) m^{DC}_{R-1}\right) && (3.35)\\ & \quad \left[R^f + \alpha^{\tau}_{R-1}\left(\widetilde{R}^e_R - R^f\right)\right]\left(R^G_R R^N_R\right)^{-1} + (1-\tau^o) ss - h_R && (3.36)\\ w^{DC}_R &= \left(w^{DC}_{R-1} + m^{DC}_{R-1}\right)\left[R^f + \alpha^{DC}_{R-1}\left(\widetilde{R}^e_R - R^f\right)\right]\left(R^G_R R^N_R\right)^{-1} && (3.37) \end{aligned}$$

and the non-negative constraint (i.e. no borrowing) in both accounts

$$w^{\tau}_{R-1} - c_{R-1} - (1-\tau^y) m^{DC}_{R-1} \geq 0, \text{ and } m^{DC}_{R-1} \geq 0 \quad (3.38)$$

Notice that the value function is changed from with one state variable, $v(w_R)$, to the value function with two state variables $\varpi\left(w^{\tau}_{R-1}, w^{DC}_{R-1}\right)$.

Follow Carroll's idea, we define two new wealth variables, namely the amount available for investment in the taxable account $a^{\tau}_{R-1} = w^{\tau}_{R-1} - c_{R-1} - (1-\tau^y) m^{DC}_{R-1}$, and the amount available for investment in the DC account $a^{DC}_{R-1} = w^{DC}_{R-1} + m^{DC}_{R-1}$. We then choose a non-negative 1-D grid to discretize $a^{\tau}_{R-1} = \{a^{\tau}_j\}^J_{j=1} \geq 0$, and do the same for $a^{DC}_{R-1} = \{a^{DC}_h\}^H_{h=1} \geq 0$.

Optimize portfolio's

The first step is to compute the optimal portfolio strategies for both accounts. The first order conditions w.r.t. α^{τ}_{R-1} and α^{DC}_{R-1} are

$$\begin{aligned} 0 &= \beta E_{R-1}\left[u'\left(w_R\left(a^{\tau}_j, a^{DC}_h\right)\right)\left(\widetilde{R}^e_R - R^f\right)\left(R^G_R R^N_R\right)^{-\gamma}\right] && (3.39)\\ 0 &= \beta E_{R-1}\left[u'\left(w_R\left(a^{\tau}_j, a^{DC}_h\right)\right)\left(\widetilde{R}^e_R - R^f\right)\left(R^G_R R^N_R\right)^{-\gamma}\right](1-\tau^o) && (3.40) \end{aligned}$$

Notice that the portfolio's are function of a^{τ} and a^{DC}. We need to determine α_{R-1}^{τ} and α_{R-1}^{DC} for each combination of $\left(a_j^{\tau}, a_h^{DC}\right)$. One special case is $\omega_{R-1}^{\tau} = \omega_{R-1}^{DC} = \omega_{R-1}$. Assuming $\omega_{R-1}^{\tau} = \omega_{R-1}^{DC} = \omega_{R-1}$ for any combination of $\{a_j^{\tau}, a_h^{DC}\}$. First, for a given vale of investable wealth in TA and TDA, we simulate the next period total wealth for certain portfolio choice α :

$$\widetilde{w}_R^{\tau}\left(a_j^{\tau}\right) = a_j^{\tau}\left[R^f + \alpha\left(\widetilde{R}_R^e - R^f\right)\right]\left(R_R^G R_R^N\right)^{-1} + (1-\tau^o)ss - h_R \quad (3.41)$$

$$\widetilde{w}_R^{DC}\left(a_h^{DC}\right) = a_h^{DC}\left[R^f + \alpha\left(\widetilde{R}_R^e - R^f\right)\right]\left(R_R^G R_R^N\right)^{-1} \quad (3.42)$$

The optimal portfolio ω_{R-1}^{*} is the one that solves the following equation based on the FOC w.r.t α

$$0 = \beta E_{R-1}\left[v'\left(\widetilde{w}_R\right)\left(\widetilde{R}_R^e - R^f\right)\left(R_R^G R_R^N\right)^{-\gamma}\right] \quad (3.43)$$

$$= \beta E_{R-1}\left[u'\left(c_R\left(\widetilde{w}_R^{\tau}, \widetilde{w}_R^{DC}\right)\right)\left(\widetilde{R}_R^e - R^f\right)\left(R_R^G R_R^N\right)^{-\gamma}\right] \quad (3.44)$$

where $v'\left(\widetilde{w}_R\right) = u'\left(c_R\left(\widetilde{w}_R^{\tau}, \widetilde{w}_R^{DC}\right)\right)$, and $c_R\left(\widetilde{w}_R^{\tau}, \widetilde{w}_R^{DC}\right)$ is obtained by interpolating the previously constructed policy function $c_{R,i,j}^{*}\left(w_{R,i,j}^{\tau}, w_{R,i,j}^{DC}\right)$.

Optimize consumption

The second step is to calculate the optimal consumption for each combination of $\{a_j^{\tau}, a_h^{DC}\}$. We know the first order conditions w.r.t. c_{R-1} is

$$u'\left(c_{R-1}\right) \quad (3.45)$$

$$= \beta E_{R-1}\left[v'\left(\widetilde{w}_R\right)\left(R^f + \alpha_{R-1}^{*}\left(\widetilde{R}_R^e - R^f\right)\right)\left(R_R^G R_R^N\right)^{-\gamma}\right] \quad (3.46)$$

$$= \beta E_{R-1}\left[u'\left(c_R\left(\widetilde{w}_R^{\tau}, \widetilde{w}_R^{DC}\right)\right)\left(R^f + \alpha_{R-1}^{*}\left(\widetilde{R}_R^e - R^f\right)\right)\left(R_R^G R_R^N\right)^{-\gamma}\right] \quad (3.47)$$

Therefore, we compute c_{R-1}^{*} directly for each underlying $\left(a_j^{\tau}, a_h^{DC}\right)$ as

$$c^*_{R-1}\left(a^\tau_j, a^{DC}_h\right) \tag{3.48}$$
$$= \left\{\beta E_{R-1}\left[u'\left(c_R\left(\widetilde{w}^\tau_R, \widetilde{w}^{DC}_R\right)\right)\left(R^f + \omega^*_{R-1}\left(\widetilde{R}^e_R - R^f\right)\right)\left(R^G_R R^N_R\right)^{-\gamma}\right]\right\}^{-1/\gamma} \tag{3.49}$$

Define a new variable called before-consumption wealth $b_{R-1} = a^\tau + c^*_{R-1} = w^\tau_{R-1} - (1-\tau^y)m^{DC}_{R-1}$. Now we can generate an endogenous 2-D grid for b_{R-1} using

$$b^*_{R-1}\left(a^\tau_j, a^{DC}_h\right) = a^\tau_j + c^*_{R-1}\left(a^\tau_j, a^{DC}_h\right). \tag{3.50}$$

In addition, we can evaluate the expected utility of the next period, for the given value of $\{a^\tau_j, a^{DC}_h\}$, as

$$EV_{R-1}\left(a^\tau_j, a^{DC}_h\right) = \beta E_{R-1}\left[\left(R^G_R R^N_R\right)^{1-\gamma} v\left(w_R\right)\right] \tag{3.51}$$
$$= \beta E_{R-1}\left[\left(R^G_R R^N_R\right)^{1-\gamma} u\left(w_R\right)\right] K\left(T-R\right) \tag{3.52}$$

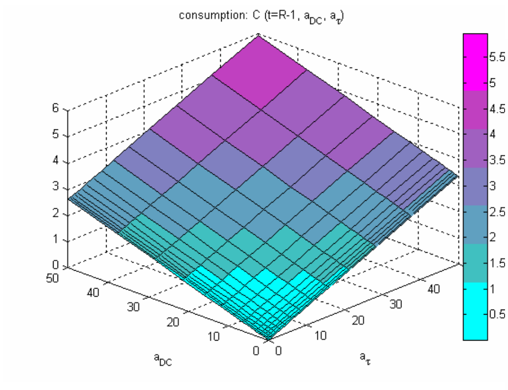


Figure: Consumption as function of investable wealth $c^*_{R-1}\left(a^\tau, a^{DC}\right)$ at time $R-1$.

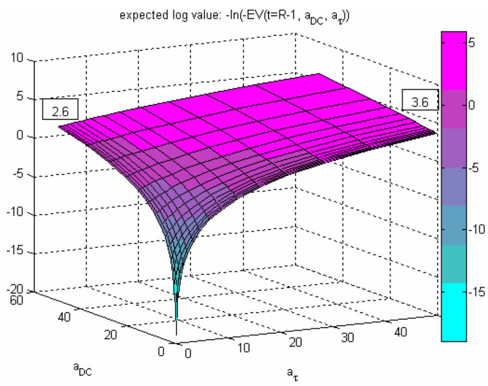


Figure: The expected value at time $R-1$, $EV_{R-1}\left(a^\tau, a^{DC}\right)$.

Optimize contribution rate

The third step is to compute the optimal contribution rate *0.2*$\geq m_{R-1} \geq 0$. It sets an upper limit on contribution rate, i.e. the contribution is no larger than 20% of gross salary income.

FOC condition w.r.t. m_{R-1} can not solve the optimal contribution rate.[11] We need to use the value function itself to search the optimal contribution rate numerically. First construct two exogenous grids for possible values of wealth $\widehat{w}^{\tau} = \{W_n\}_{n=1}^{N}$ and $\widehat{w}^{DC} = \{W_n\}_{n=1}^{N}$. Then, we construct a grid of possible contribution rates. We start with the case without employer match. The grid of possible contribution rates is denoted by $\widehat{m}_{R-1} = \{m_i\}_{i=1}^{M} \subset [0, m^{\max}]$, with $m^{\max} = \min\left(1, \widehat{w}^{\tau}/(1-\tau^{y})\right)$. These implies a set of before-consumption wealth $\widehat{b}$ and investable wealth $\widehat{a}^{DC}$ for any given combination of $\{\widehat{w}^{\tau}, \widehat{w}^{DC}, \{m_i\}_{i=1}^{M}\}$, as

$$\widehat{b}_i = \widehat{w}^{\tau} - (1-\tau^{y})\widehat{m}_i > 0 \tag{3.53}$$
$$\widehat{a}_i^{DC} = \widehat{w}^{DC} + \widehat{m}_i \tag{3.54}$$

With employer match, e.g. $m_t^{DC} = \min\left(m_t + \min\left(m_t, 6\%\right), 20\%\right)$, the grid of possible contribution rates remains the same, but the implied $\widehat{b}$ and $\widehat{a}^{DC}$ become

$$\widehat{b}_i = \widehat{w}^{\tau} - (1-\tau^{y})\widehat{m}_i > 0 \tag{3.55}$$
$$\widehat{a}_i^{DC} = \widehat{w}^{DC} + \widehat{m}_i^{DC} \tag{3.56}$$

If $\widehat{b}_i$ and $\widehat{a}_i^{DC}$ are known, then with the help of the interpolation relation

[11] FOC w.r.t. m_{R-1}^{DC} is

$$(1-\tau^{y})\beta E_{R-1}\left[\left(R_R^G R_R^N\right)^{-\gamma} v'\left(w_R\right)\widetilde{R}_R^{P,\tau}\right] = (1-\tau^{o})\,\beta E_{R-1}\left[\left(R_R^G R_R^N\right)^{-\gamma} v'\left(w_R\right)\widetilde{R}_R^{P,DC}\right]$$

which clearly doesn't hold in general. Since $\tau^{y} > \tau^{o}$, this FOC implies marginal cost of contribution < martinal benefit of contribution, therefore m_{R-1}^{DC} should take some maximum value (if exists). However the danger of doing so is that it implies w_{R-1}^{τ} might be unlimited. Due to this reason, we need to resort to value function to find the optimal m_{R-1}^{DC} for any given level of $\left\{w^{\tau}, w^{DC}\right\}$.

$b^*_{R-1}\left(a^\tau, a^{DC}\right)$ we can back out the corresponding private wealth $\widehat{a}^\tau_i$. It then leads to the optimal consumption $\widehat{c}^*_i = \widehat{b}_i - \widehat{a}^\tau_i$, for any given set of $\{\widehat{w}^\tau, \widehat{w}^{DC}, m_i\}$. We denote the implied consumption as $c^*_i\left(\widehat{w}^\tau, \widehat{w}^{DC}, m_i\right)$. Furthermore, we can compute the expected value of the next period $\widehat{EV}_i\left(\widehat{a}^\tau_i, \widehat{a}^{DC}_i\right)$ based on relation $EV_{R-1}\left(a^\tau, a^{DC}\right)$. Finally, we can evaluate the trade-off between the consumption and the contribution using the recursive objective function as

$$\varpi_{R-1}(\widehat{w}^\tau_{R-1}, \widehat{w}^{DC}_{R-1}) \equiv \max_{\{m_i\}_{i=1}^M} \frac{\left(\widehat{c}^*_i\right)^{1-\gamma}}{1-\gamma} + EV_{R-1}\left(\widehat{a}^\tau_i, \widehat{a}^{DC}_i\right) \tag{3.57}$$

The optimal contribution rate is the one that maximize the above expression for given values of $(\widehat{w}^\tau_{R-1}, \widehat{w}^{DC}_{R-1})$. Let's denote it as $m^*(\widehat{w}^\tau, \widehat{w}^{DC})$. As a by-product, we also get the corresponding consumption policy $c^*(w^\tau, w^{DC}) = c^*\left(\widehat{w}^\tau, \widehat{w}^{DC}, m^*(\widehat{w}^\tau, \widehat{w}^{DC})\right)$.

Euler equation

Finally, the Envelope theorem tells us

$$\begin{aligned}
& \varpi'_\tau\left(w^\tau_{R-1}, w^{DC}_{R-1}\right) & (3.58)\\
= \; & \beta E_{R-1}\left[\left(R^G_R R^N_R\right)^{-\gamma} v'\left(w_R\right)\left[R^f + \alpha_{R-1}\left(\widetilde{R}^e_R - R^f\right)\right]\right] & (3.59)\\
= \; & u'\left(c^*_{R-1}\left(w^\tau_{R-1}, w^{DC}_{R-1}\right)\right) & (3.60)\\
& \varpi'_{DC}\left(w^\tau_{R-1}, w^{DC}_{R-1}\right) & (3.61)\\
= \; & \left(1-\tau^o\right)\beta E_{R-1}\left[\left(R^G_R R^N_R\right)^{-\gamma} v'\left(w_R\right)\left[R^f + \alpha_{R-1}\left(\widetilde{R}^e_R - R^f\right)\right]\right] & (3.62)\\
= \; & \left(1-\tau^o\right) u'\left(c^*_{R-1}\left(w^\tau_{R-1}, w^{DC}_{R-1}\right)\right) & (3.63)
\end{aligned}$$

where ϖ'_τ and ϖ'_{DC} denote the partial differentials.

Optimal policies for $t = R-1$

For the last year of working, $t = R-1$, the optimal contribution rate is largely 100% of the labor income, except for when TA wealth (w^τ) is very limited, but the TDA wealth (w^{DC}) is abundant. The optimal consumption increases with TA and

TDA wealth in general. The special feature for the optimal consumption is that there are kinks due to the liquidity constraint, i.e. individual can not consume more than what they have in the taxable account. Similarly there are also kinks for the value function. Except for this, the value function increases with both wealth accounts.

Repeat the procedure

Now we are able to proceed to $t = R-2, R-3, ...1$, by repeating the similar procedure as for $t = R-1$. To generate the average pattern of life cycle portfolio holding, as in Figure 1, we simulate the model from time 1 to T for 10,000 scenario's, and take the average over all simulated scenario's.

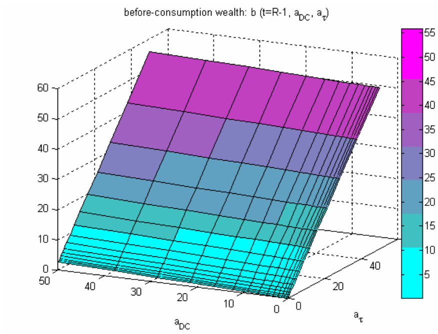


Figure: The implied before-consumption wealth b as function of investable wealth $b_{R-1}\left(a^{\tau}, a^{DC}\right)$ at time $R-1$.

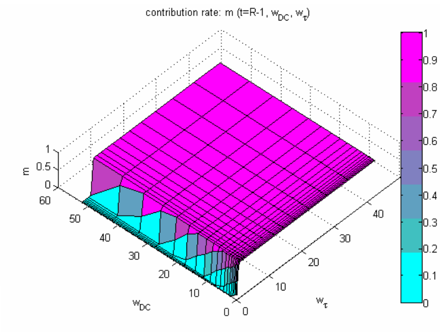


Figure: The optimal contribution rate $m^{*}_{R-1}\left(w^{\tau}, w^{DC}\right)$ at time $R-1$.

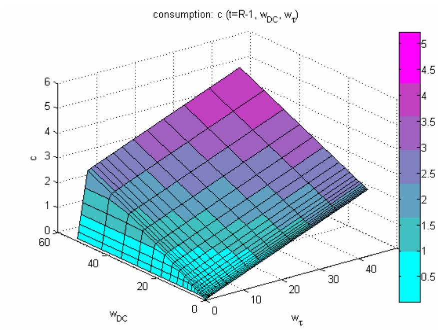


Figure: The optimal consumption policy $c^{*}_{R-1}\left(w^{\tau}, w^{DC}\right)$.

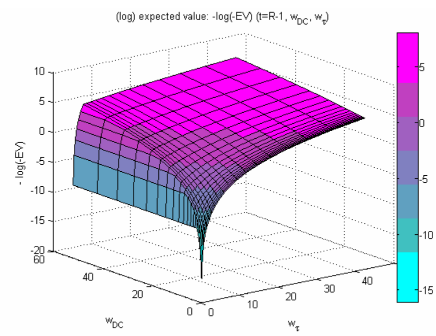


Figure: (log) value function -ln$(-\varpi(w^{\tau}, w^{DC}))$ at time R-1.

3.8 Appendix B: Solution method of Section 4

In Default #1, the contribution rate $m \geq 0$ and portfolio allocations $0.2 \geq \alpha \geq 0$ are chosen at the beginning of the career ($t = 0$), and are fixed throughout life time. Notice that we the portfolio allocation for both accounts are assumed to be the same constant mix through out life time, $\alpha^{\tau} = \alpha^{DC} = \alpha$. Therefore, the value functions, for given the chosen level of m and α, are denoted as $\varpi\left(w_t^{\tau}, w_t^{DC} \mid \alpha, m\right)$ for working period, and $v\left(w_t \mid \alpha, m\right)$ for retirement period. We solve the model numerically using dynamic programming.

For the final period, the optimal consumption policy is to consume everything $c_T^* = w_T$. The corresponding value function is given by $v\left(w_T \mid \alpha, m\right) = \frac{\left(c_T^*\right)^{1-\gamma}}{1-\gamma}$. Then we proceed to time $t = T-1$. During the retirement period ($R \leq t \leq T-1$), the normalized value function (in recursive form) and wealth process are

$$v\left(w_t \mid \alpha, m\right) = \frac{c_t^{1-\gamma}}{1-\gamma} + \beta E_t\left[\left(R_{t+1}^G\right)^{1-\gamma} v\left(w_{t+1} \mid \omega, m\right)\right] \tag{3.64}$$

subject to the budget dynamics

$$w_{t+1} = \left(w_t - c_t\right)\widetilde{R}_{t+1}^{P,\omega}\left(R_{t+1}^G\right)^{-1} + (1-\tau)\,ss_{t+1} - h_{t+1} - med_{t+1} \tag{3.65}$$

where we use a shorthand notation $\widetilde{R}_{t+1}^{P,\omega} = R^f + \alpha\left(\widetilde{R}_{t+1}^e - R^f\right)$ for the portfolio returns.

For any given value of m and α, we only need to optimize the consumption choice c_t.

The FOC w.r.t. c_t is

$$u'\left(c_t\right) = \beta E_t\left[v'\left(w_{t+1} \mid \alpha, m\right)\widetilde{R}_{t+1}^{P,\omega}\left(\widetilde{R}_{t+1}^G\right)^{-\gamma}\right] \tag{3.66}$$

The Envelope theorem gives

$$v'\left(w_t \mid \alpha, m\right) = \beta E_t\left[v'\left(w_{t+1} \mid \alpha, m\right)\widetilde{R}_{t+1}^{P,\omega}\left(R_{t+1}^G\right)^{-\gamma}\right] = u'\left(c_t\right) \tag{3.67}$$

So, pushing one period ahead, we have $v'\left(w_{t+1} \mid \alpha, m\right) = u'\left(c_{t+1}\right)$.

Define a new variable, $a_t = w_t - c_t$, as the after-consumption-wealth. Construct a grid of $a_t = \{a_j\}_{j=1}^J$. Now we solve for the optimal consumption and portfolio policies for each given a_j.

$$c_t = I_u\left(\beta E_t\left[v'\left(w_{t+1}\mid\alpha,m\right)\widetilde{R}^{P,\omega}_{t+1}\left(\widetilde{R}^G_{t+1}\right)^{-\gamma}\right]\right) \tag{3.68}$$

$$= I_u\left(\beta E_t\left[u'\left(c^*_{t+1}[w_{t+1}]\right)\widetilde{R}^{P,\omega}_{t+1}\left(\widetilde{R}^G_{t+1}\right)^{-\gamma}\right]\right) \tag{3.69}$$

Following the EGM, the optimal wealth process is endogenously determined by $w^*_t = c^*_t\left(a_t\right) + a_t$, for a given value of $\{m,\,\alpha\}$.

At time $t = R$, we split the single wealth variable w^*_R into two wealth variables w^τ_R and w^{DC}_R, and obtain the corresponding policy function $c^*_R\left(w^\tau_R, w^{DC}_R\right)$, as done in Appendix A.

During the working period ($0 \le t \le R-1$), the normalized value function (in recursive form) and wealth process are

$$\varpi\left(w^\tau_t, w^{DC}_t\mid\alpha,m\right) = \frac{c_t^{1-\gamma}}{1-\gamma} + \beta E_t\left[\left(R^G_{t+1}\right)^{1-\gamma}\varpi\left(w^\tau_{t+1}, w^{DC}_{t+1}\mid\omega,m\right)\right]$$

s.t. the wealth dynamics and no borrowing constraint as

$$w^\tau_{t+1} = \left(w^\tau_t - c_t - (1-\tau^y)m\right)\widetilde{R}^{P,\omega}_{t+1}\left(R^G_{t+1}R^N_{t+1}\right)^{-1} \tag{3.70}$$

$$+(1-\tau^y) - h_{t+1} - med_{t+1} \tag{3.71}$$

$$w^{DC}_{t+1} = \left(w^{DC}_t + m^{DC}\right)\widetilde{R}^{P,\omega}_{t+1}\left(R^G_{t+1}R^N_{t+1}\right)^{-1} \tag{3.72}$$

$$0 \le w^\tau_t - c_t - (1-\tau^y)m \tag{3.73}$$

Construct two investable wealth grids $a^\tau_t \equiv \left(w^\tau_t - c_t - (1-\tau^y)m\right) = \{a^\tau_j\}^J_{j=1} \ge 0$, and $a^{DC}_t \equiv \left(w^{DC}_t + m^{DC}\right) = \{a^{DC}_h\}^H_{h=1} \ge 0$. Then calculate the optimal consumption for each combination of $\{a^\tau_j, a^{DC}_h\}$. We know the first order conditions w.r.t. c_t is

$$u'\left(c_t\right) = \beta E_t\left[v'\left(\widetilde{w}_{t+1}\right)\widetilde{R}^{P,\omega}_{t+1}\left(R^G_{t+1}R^N_{t+1}\right)^{-\gamma}\right] \tag{3.74}$$

$$= \beta E_{R-1}\left[u'\left(c_{t+1}\left(\widetilde{w}^\tau_{t+1}, \widetilde{w}^{DC}_{t+1}\right)\right)\widetilde{R}^{P,\omega}_{t+1}\left(R^G_{t+1}R^N_{t+1}\right)^{-\gamma}\right] \tag{3.75}$$

Therefore, we compute c_t^* directly for each underlying $\left(a_j^\tau, a_h^{DC}\right)$ as

$$c_t^*\left(a_j^\tau, a_h^{DC}\right) = \left\{\beta E_t\left[u'\left(c_{t+1}\left(\widetilde{w}_{t+1}^\tau, \widetilde{w}_{t+1}^{DC}\right)\right)\widetilde{R}_{t+1}^{P,\omega}\left(R_{t+1}^G R_{t+1}^N\right)^{-\gamma}\right]\right\}^{-1/\gamma} \tag{3.76}$$

Following the EGM, the optimal wealth process is endogenously determined by $w_t^{*,\tau} = c_t^*\left(a_t^\tau, a^{DC}\right) + a_t^\tau + (1-\tau^y)m$ and $w_t^{*,DC} = a^{DC} - m^{DC}$, for a given value of $\{m,\, \alpha\}$.

Repeat the above procedure for all values of $\{m,\, \alpha\}$. Finally, the optimal $\{\alpha^*, m^*\}$ are the ones that maximize the value function

$$\{\alpha^*, m^*\} = \arg\max \varpi\left(w_0^\tau, w_0^{DC} \mid \alpha, m\right)$$

For Defaults 2, 3 and 4, a similar procedure is applied when determining the opimal shape of the contribution and portfolio rules.

3.9 Tables and Figures

Table 2: **Default #1, constant contribution rate and constant portfolio**, (**no** employer matching).

This table presents the welfare (CEC) under given flat contribution rates and several different portfolio weights in equities. ($CEC^{TDATA} = 0.6885$)

contrib.	Optimal portfolio α^*			50% in equities		100% risk free ($\alpha = 0$)	
m	α^*	CEC	$\frac{CEC}{CEC^{TDATA}}$	CEC	$\frac{CEC}{CEC^{TDATA}}$	CEC	$\frac{CEC}{CEC^{TDATA}}$
0%	80%	0.675	98%	0.672	97.6%	0.661	96%
1%	80%	0.671	97.5%	0.67	97.2%	0.6534	95%
2%	80%	0.666	96.8%	0.664	96.5%	0.649	94.2%
3%	82.5%	0.661	96%	0.659	95.6%	0.644	93.5%
4%	**82.5%**	**0.654**	**95%**	**0.653**	**94.8%**	**0.638**	**92.7%**
5%	85%	0.648	94.2%	0.646	94%	0.633	91.9%

Table 3: **Default #1, constant contribution rate and constant portfolio**, (**with** employer matching).

This table presents the welfare (CEC) under given flat contribution rates and several different portfolio weights in equities. ($CEC^{TDATA} = 0.702$).

contrib.	Optimal portfolio α^*			50% in equities		100% risk free ($\alpha = 0$)	
m	α^*	CEC	$\frac{CEC}{CEC^{TDATA}}$	CEC	$\frac{CEC}{CEC^{TDATA}}$	CEC	$\frac{CEC}{CEC^{TDATA}}$
0%	80%	0.65	96.1%	0.672	95.7%	0.661	94.1%
1%	80%	0.676	96.3%	0.674	96%	0.658	93.7%
2%	80%	0.674	96%	0.672	95.7%	0.656	93.5%
3%	82.5%	0.670	95.4%	0.668	95.2%	0.654	93.2%
4%	**82.5%**	**0.664**	**94.6%**	**0.663**	**94.4%**	**0.651**	**92.8%**
5%	85%	0.656	93.5%	0.656	93.4%	0.647	92.1%

Table 4: **Welfare comparison across default designs** (without employer matching)

Designs	default specification		Welfare measure	
	portfolio	contribution	CEC	$\frac{CEC}{CEC^{TDATA}}$
Fully Optimal TDA & TA	α_t^*	m_t^*	0.688	100%
Current Default	$\alpha = 0$	$m = 4\%$	0.638	92.7%
D#1 (with better portfolio)	$\alpha = 50\%$	$m = 4\%$	0.653	94.8%
D#1 (with best portfolio)	$\alpha = 82.5\%$	$m = 4\%$	0.654	95%
D#2 (age-dep. portfolio)	α age-dep.	$m = 4\%$	0.655	95.2%
D#3 (age-dep. contribution)	$\alpha = 0$	m age-dep.	0.668	97.1%
D#3 (age-dep. contrib., best portf.)	$\alpha^* = 85\%$	m age-dep.	0.683	99.2%
D#4 (age-dep contrib, age-dep portf)	α age-dep.	m age-dep.	0.684	99.4%

Table 5: **Welfare comparison across default designs** (**with** employer matching)

Designs	default specification		Welfare measure	
	portfolio	contribution	CEC	$\frac{CEC}{CEC^{TDATA}}$
Fully Optimal TDA & TA	α_t^*	m_t^*	0.702	100%
Current Default	$\alpha = 0$	$m = 4\%$	0.651	92.8%
D#1 (with better portfolio)	$\alpha = 50\%$	$m = 4\%$	0.663	94.4%
D#1 (with best portfolio)	$\alpha = 82.5\%$	$m = 4\%$	0.664	94.6%
D#2 (age-dep. portfolio)	α age-dep.	$m = 4\%$	0.668	95.1%
D#3 (age-dep. contribution)	$\alpha = 0$	m age-dep.	0.681	97%
D#3 (age-dep. contrib., best portf.)	$\alpha^* = 85\%$	m age-dep.	0.695	99%
D#4 (age-dep. contrib age-dep portf)	α age-dep.	m age-dep.	0.697	99.2%

Figure 1: **Income profiles.** (a) Left panel shows the life cycle profile of the age-dependent idiosyncratic component N_{t-t_0}; (b) Right panel shows the aggregate wage component, G_t.

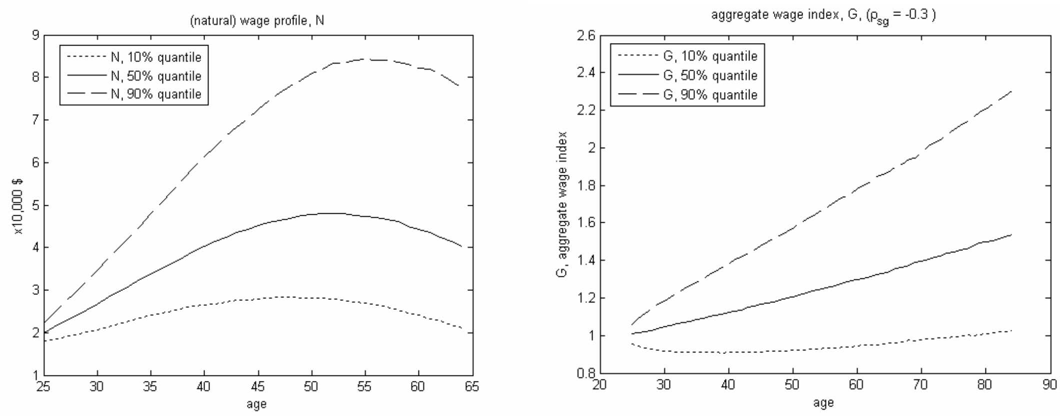


Figure 2: **Housing and medical expenditures**. (a) Left panel shows the mean and one simulation of the ratio of housing expenditure to income over life cycle; (b) Right panel shows the mean and one simulation of the ratio of out-of-pocket medical expenditure to income over life cycle.

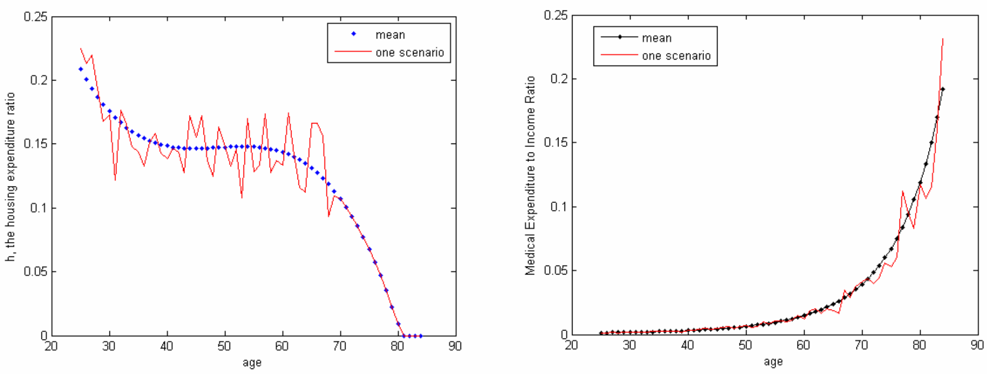


Figure 3: **Optimal portfolio choice and consumption profiles in TA/TDA.** (left panel) Life-cycle portfolio choice assuming the identical asset allocation in Taxable and Tax-deferred DC accounts (5%, 50%, 95% quantiles); (right panel) Life-cycle consumption profile (5%, 50%, 95% quantiles).

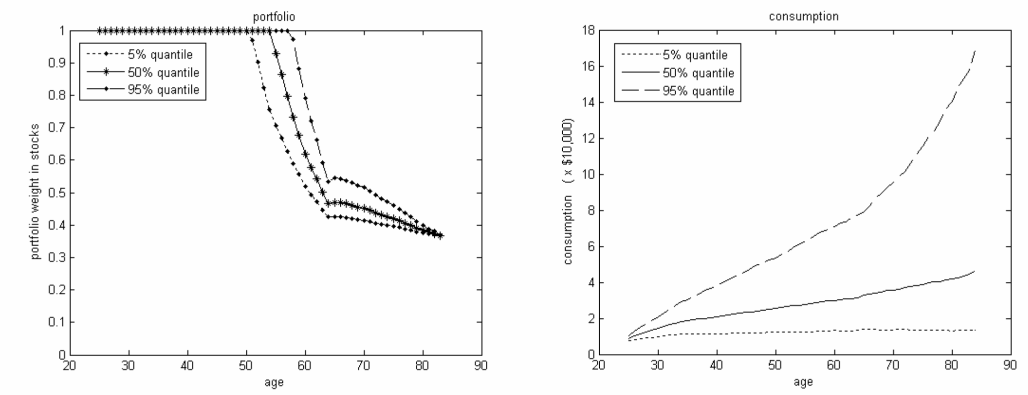


Figure 4: **Optimal contribution rate and wealth accumulation in TA/TDA.** (left panel) life-cycle contribution rate m_t (5%, 50%, 95% quantiles), and (right panel) wealth accumulation in Taxable and Tax-deferred DC accounts (5%, 50%, 95% quantiles).

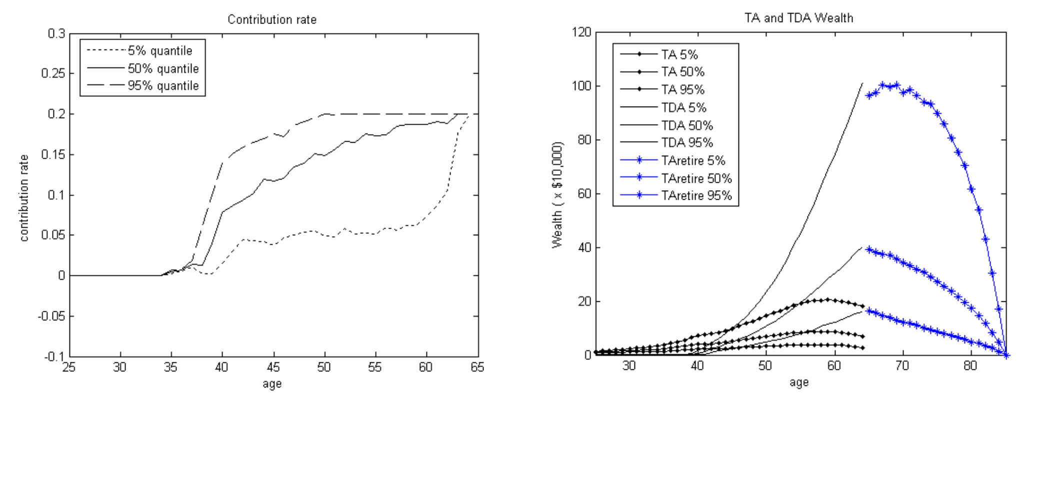


Figure 5: **Optimal portfolio choice and consumption profiles in TA/TDA with employer matching.** (left panel) Life-cycle portfolio choice assuming the identical asset allocation in Taxable and Tax-deferred DC accounts, (5%, 50% and 95% quantiles); (right panel) Life-cycle consumption **with employer match**.

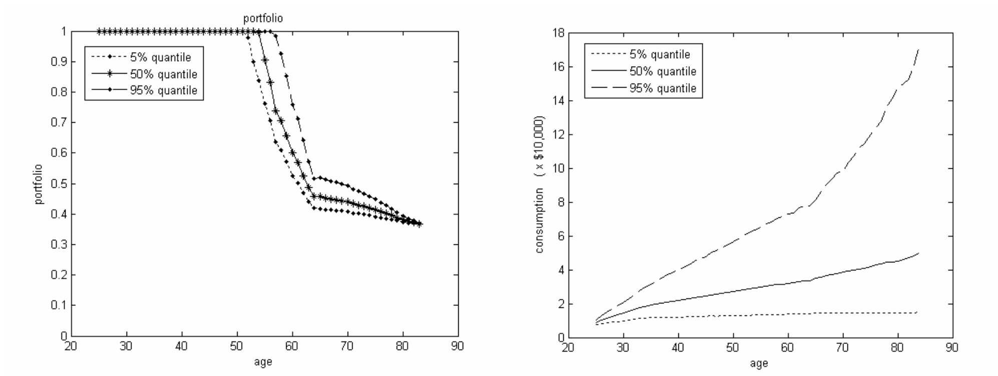


Figure 6: **Optimal contribution rate and wealth accumulation in TA/TDA with employer matching.** (left panel) life-cycle contribution rate m_t with employer match (5%, 50% and 95% quantiles), and (right) asset accumulation in Taxable and Tax-deferred DC accounts with employer match.

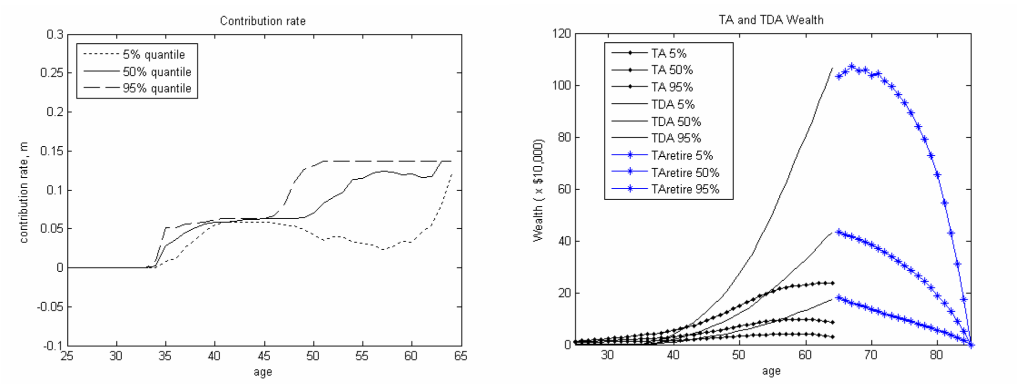


Figure 7: **Age-dependent portfolio rule in Default #2**. The optimized age-dependent portfolio rule for given contribution rate of 4%, as specified in Default #2.

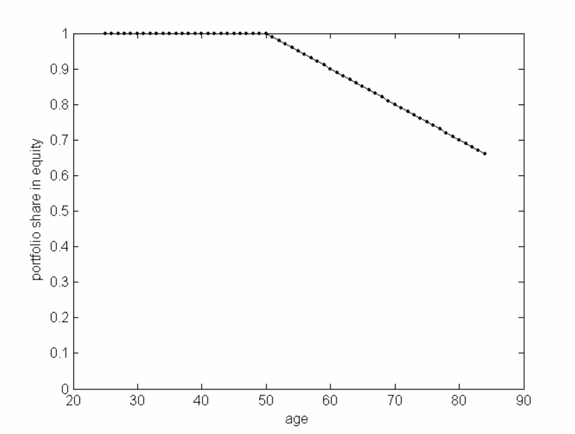


Figure 8: **Consumption and wealth accumulation in Default #2.** (left) Life-cycle consumption, and (right) asset accumulation in Taxable and Tax-deferred DC accounts, as specified in Default #2 (5%, 50% and 95% quantiles).

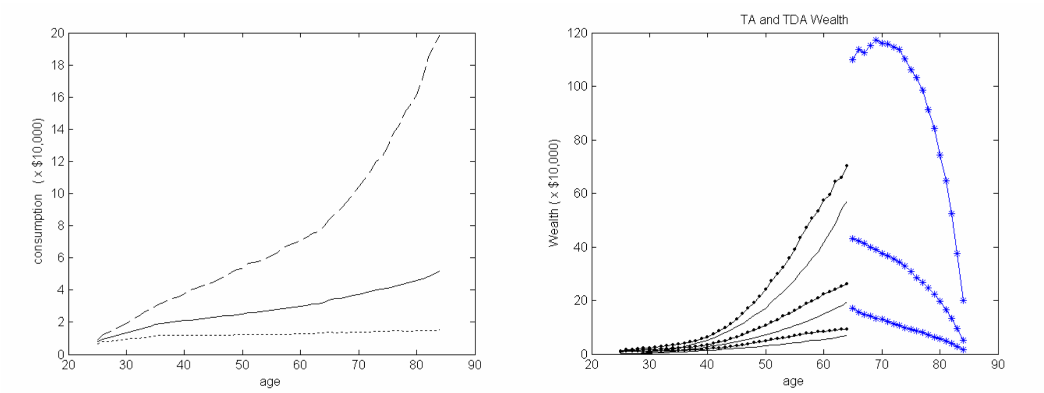


Figure 9: **The optimized age-dependent contribution rule for given portfolio choice of 85%, as specified in Default #3.**

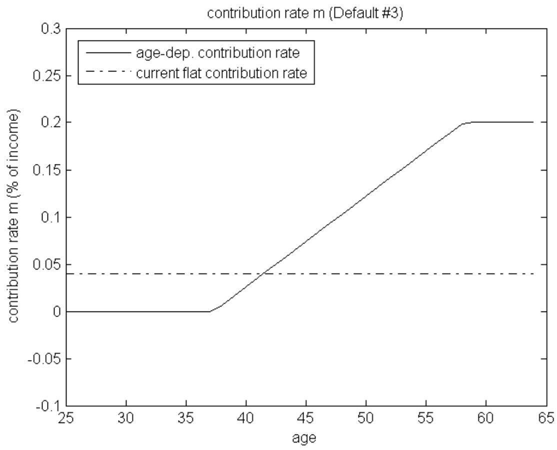


Figure 10: **Consumption and wealth accumulation in Default #3.** (left) Life-cycle consumption, and (right) asset accumulation in Taxable and Tax-deferred DC accounts, as specified in Default #3 (5%, 50% and 95% quantiles).

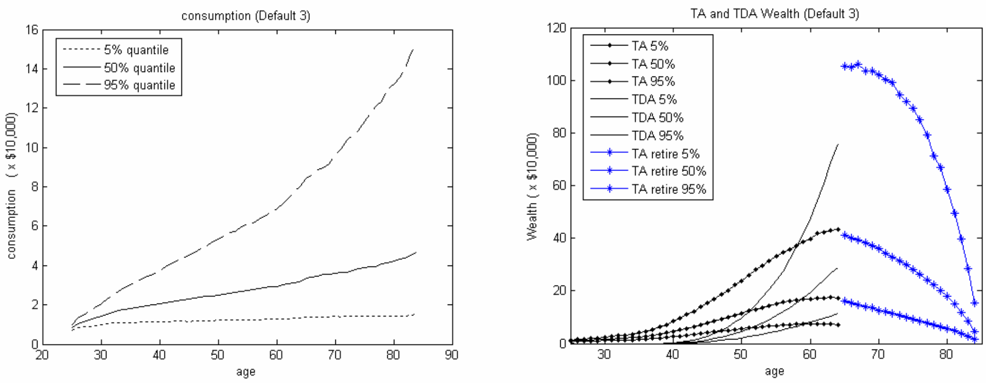


Figure 11: **The optimized age-dependent contribution rate and portfolio rule in Default #4.** The age-dependent contribution rate (left panel) and age-dependent portfolio rule (right panel) in Default #4 (**no** employer matching).

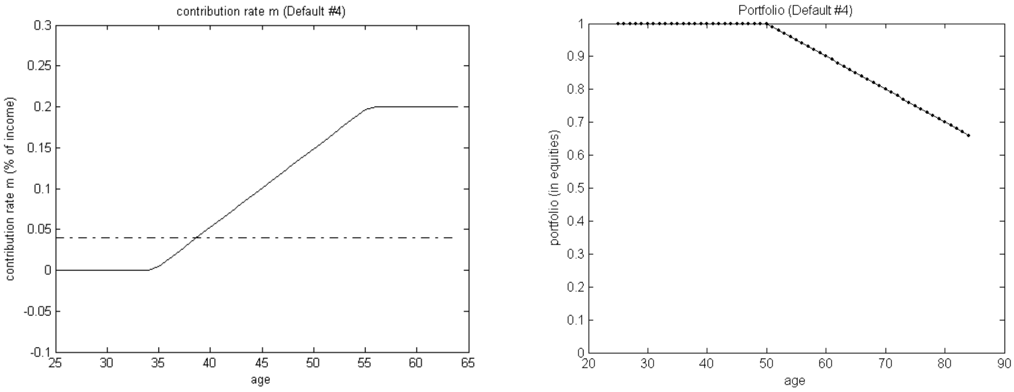


Figure 12: **The optimized age-dependent contribution rate and portfolio rule in Default #4 with employer matching.** The age-dependent contribution rate (left panel) and age-dependent portfolio rule (right panel) in Default #4 (**with** employer matching).

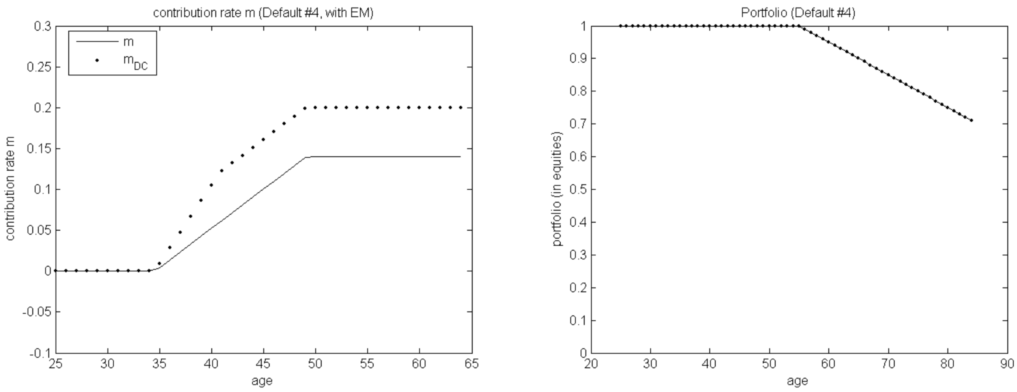


Figure 13: **Sensitivity analysis: Contribution rule for different constant asset mix.** The optimal age-dependent contribution rate rules (default #3) for different asset mix ($\alpha = 0\%$ and 85% respectively), based on the baseline preference assumptions ($\gamma = 5$, $\beta = 0.97$).

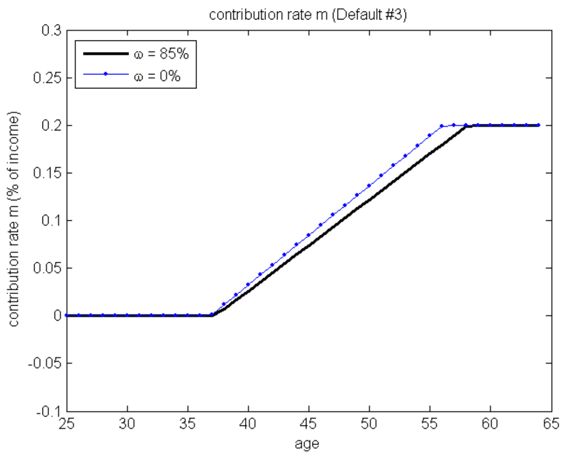


Figure 14: **Optimal contribution and portfolio choice in TA/TDA for different risk aversion.** The age profiles of the optimal contribution rates (left panel), optimal portfolio choices (right panel) of a $\gamma = 8$ individual in taxable and tax-deferred DC account, without employer matching.

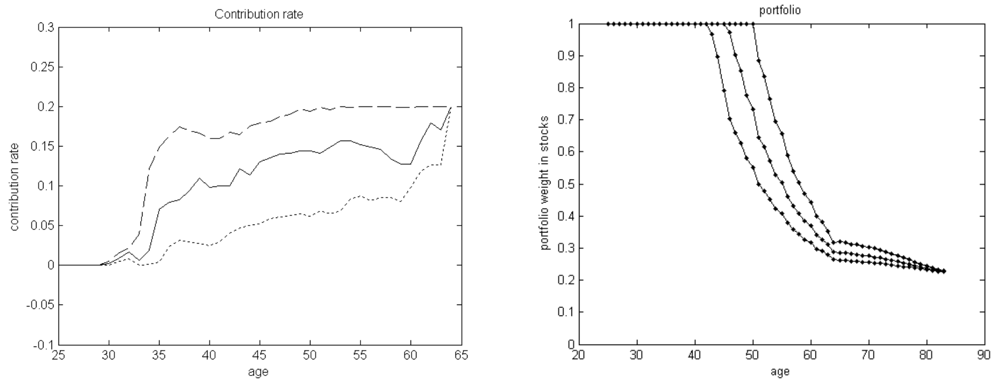


Figure 15: **Optimal contribution and portfolio choice in TA/TDA for different risk aversion.** (left panel) The life cycle consumption and (right panel) asset accumulation for a $\gamma = 8$ individual in taxable and tax-deferred DC account, without employer matching.

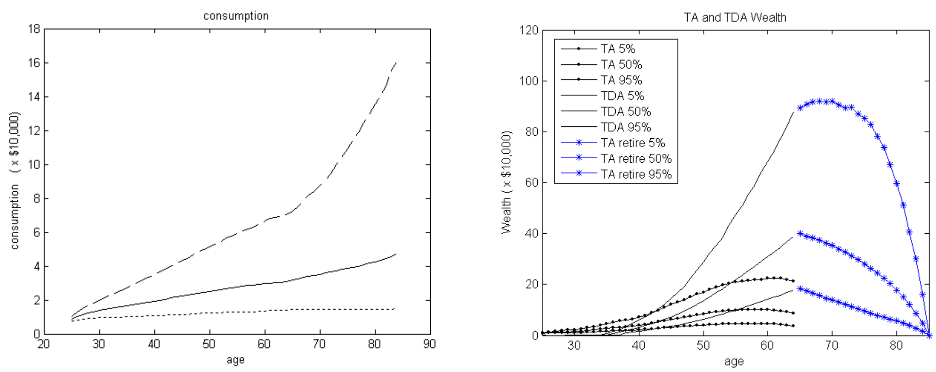


Figure 16: **The optimal age-dependent contribution rate rules (default #3) for individuals with different risk aversion ($\gamma = 5$ and 8 respectively).**

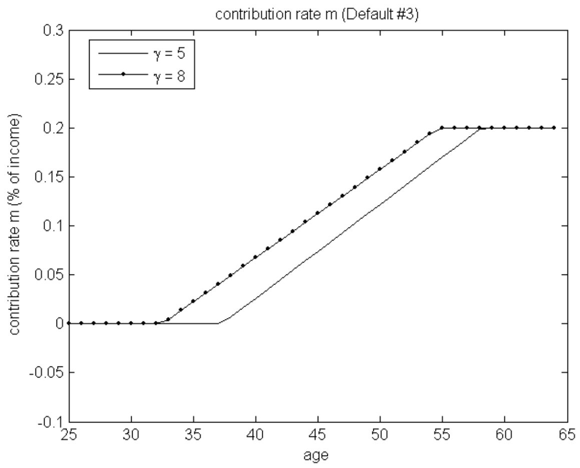


Figure 17: **Optimal portfolio and consumption profiles for the case with flat wage earnings.** (left panel) Life-cycle portfolio choice assuming the identical asset allocation in Taxable and Tax-deferred DC accounts (5%, 50% and 95% quantiles); (right panel) Life-cycle consumption profile.

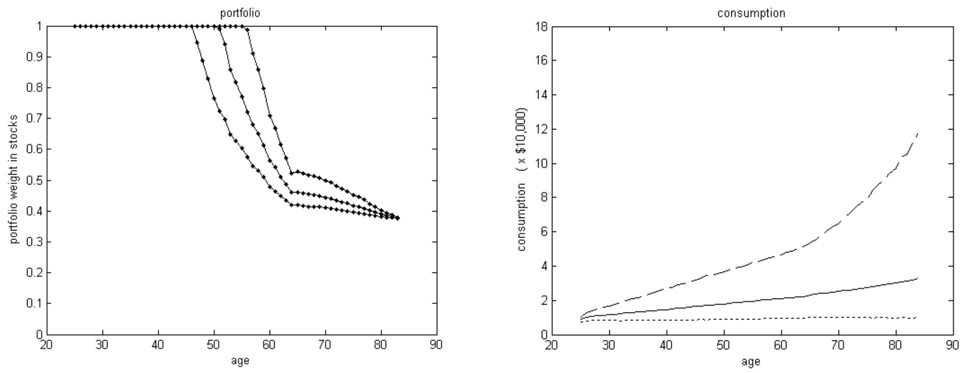


Figure 18: **Optimal contribution and wealth profiles for the case with flat wage earnings.** (left panel) life-cycle contribution rate m_t (5%, 50% and 95% quantiles), and (right panel) wealth accumulation in Taxable and Tax-deferred DC accounts.

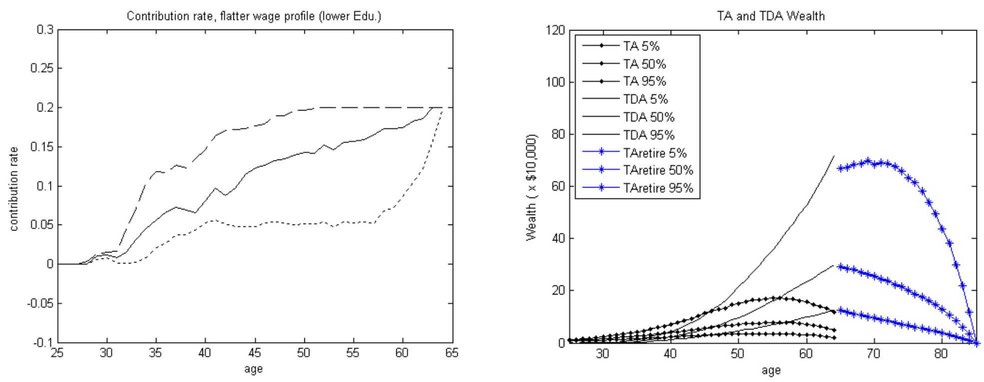
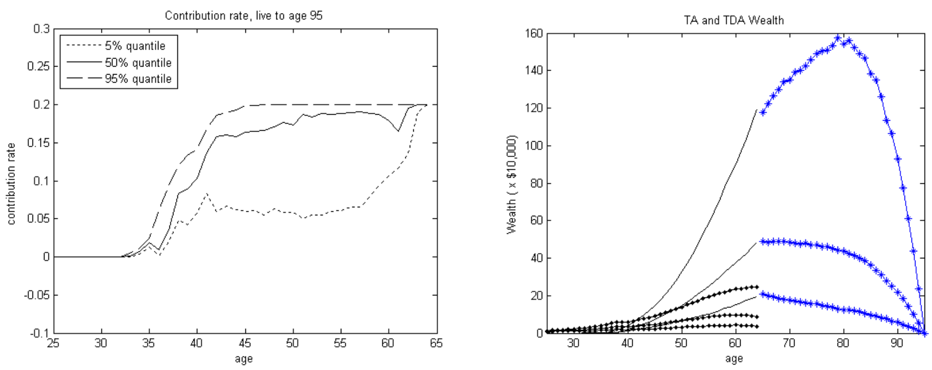


Figure 19: **Optimal contribution and wealth profiles for the case with longer life span.** (left panel) life-cycle contribution rate m_t (5%, 50% and 95% quantiles), and (right panel) wealth accumulation in Taxable and Tax-deferred DC accounts (5%, 50% and 95% quantiles).



Chapter 4

Longevity Risk Pricing

This chapter is based on Cui (2008a).

4.1 Introduction

Longevity risks, i.e., unexpected improvements in life expectancies, impose a challenge on pension plans and insurance companies because small unexpected improvement in life expectancies may lead to severe solvency issues for these annuity providers. Longevity-linked securities are designed to pay out more when a selected population group lives longer than originally expected. They are attractive securities to financial markets because, on one hand, they are desirable assets for annuity providers to hedge their longevity risks, and on the other hand, investors may find these securities attractive for the benefits of diversification provided that the risk premia are set appropriately. Moreover, financial markets may provide a more efficient risk allocation than the traditional insurance markets. Although academic researchers, policy makers and practitioners have talked about it for years, longevity-linked securities are not traded in financial markets due to the pricing difficulty. This paper therefore proposes a new method to price the longevity risk premia in order to tackle the pricing obstacles of the innovative longevity-linked securities.

This paper contributes to the literature by quantifying the longevity risk premia in various longevity-linked securities (bonds, swaps, caps and floors), applying the equivalent utility pricing principle. Based on the equivalent utility pricing principle, we obtain a minimum risk premium required by the longevity insurance seller and a maximum acceptable risk premium by the longevity insurance buyer. These upper and lower bounds indicate a price range for negotiation between the sellers and the buyers. The four main advantages of our methodology are: i) the suitability for incomplete market pricing, ii) a narrow range of the risk premia, iii) the consistency with other financial market risk premia (like inflation risk premium) and iv) its flexibility in handling different payoff structures, basis risk and natural hedging possibilities.

In practice, life insurers, also pension funds, claim that their annuity businesses are losing money due to the unexpected longevity improvements over years. In the past centuries remarkable improvements in human life expectancy have been observed. The uncertainties about the further improvements of human life expectancy are referred to as the longevity risks. Oeppon and Vaupel (2002, in *Science*) report striking evidences that the record life expectancy has been rising nearly three months per year in the past 160 years, and the asserted ceilings on life expectancy were surpassed repeatedly in the past century. In fact, the future improvements of life expectancy are difficult to be predicted accurately.[1] The general opinion from the experts tends to be that the trend of longevity improvements is certain, but deviations to both sides are possible.

Reinsurance contracts do exist, but the capacity of reinsurance is limited (OECD (2005)). By a reinsurance contract, the longevity risk is concentrated on one large reinsurance company. The reinsurance approach works best when the underlying risks are diversifiable. However, the fact that longevity risk is a systematic risk weakens the diversification principle that the reinsurance requires. Longevity risk cannot be reduced by diversification or increasing the size of the pool.[2]

Alternatively, the longevity risk could be transferred to financial markets, also known as securitization. By transferring longevity risk to financial markets, the

[1] Brown and Orszag (2006) discuss the difficulties in making an accurate mortality projection.

[2] We must differentiate mortality risk from longevity risk. Mortality risk refers to the uncertainty about individual death events when the life expectancy is known. Therefore, mortality risk is a micro risk, and can be diversified by increasing the size of the pool.

risk is distributed among a (large) number of market participants who can shoulder the risk better, i.e., achieving more efficient risk allocation. Longevity-linked securities are one of the current financial market innovations. The first longevity bond was announced by the EIB and BNP[3] Paribas in November 2004, but it has been under-subscribed, and withdrawn for redesign in late 2005.[4] The EIB/BNP survivor bond is a coupon-based bond, with floating annual coupons linked to a cohort survivor index. The problem with this issuance is that there is no clear view on how longevity risk should be charged. The EIB/BNP survivor bond required a longevity risk premium of 20 basis points, which was regarded as too high for some annuity providers. The EIB/BNP bond, although linked to the British survivor index, was also marketed among Dutch pension funds. However, it was not clear for the Dutch pension funds whether the 20 basis points was a good deal or not.

The origin of the pricing difficulty lies in the fact that the financial market is incomplete when longevity securities are not traded. Therefore, the goal of this paper is to provide a pricing framework suitable for pricing longevity risks in incomplete market setting. Based on the equivalent utility pricing principle, our method obtains the minimum risk premium required by the longevity insurance seller and the maximum acceptable risk premium by the longevity insurance buyer. We find that the size of the risk premium depends on the payoff structure of the security, the financial strength of the seller and the buyer and the availability of the natural hedge. We show that different payoff structures and maturities may lead to different risk premia because the market is incomplete. We also show that financially stronger issuer may require a lower risk premium. Furthermore, the risk premium could be reduced by distributing the longevity risk among more market participants. The market calls for more longevity bond issuers in order to achieve more efficient allocation and reduce the longevity risk premium. One

[3]EIB/BNP stands for the European Investment Bank (EIB) and Banque Nationale de Paris (BNP).

[4]Blake, Cairns and Dowd (2006) address the associated obstacles in current market development of longevity-linked securities. The obstacles are categorized into 'design issues' (regarding the payoff structure, maturity, choice of survivor index, nominal or real payments, etc.), 'pricing issues' and 'institutional issues'. As to the pricing issues, the authors comment that "even if the (survivor) bond provides a perfect hedge, there will be uncertainty over what the right price to pay or charge should be."

important implication for the market development of longevity-linked securities is that multiple sellers, instead of a single seller, are required.

In this paper, the longevity risk is modeled as proposed by Lee and Carter (1992), and estimated according to the U.K. and the Dutch mortality data. However our pricing methodology is quite general. Other stochastic mortality models are also suitable for our pricing framework.

Recently, a few approaches to price longevity risk were proposed in the literature. Friedberg and Webb (2005) apply Capital Asset Pricing Model (CAPM) and Consumption-based Capital Asset Pricing Model (CCAPM) to estimate the longevity premium. Their result based on CAPM leads to a risk premium of 75 basis points, with confidence interval ranging from -75 to 230 basis points, due to inaccuracy in estimating the beta. Their result based on CCAPM is merely two basis points, due to the low variation in consumption data. The discrepancy between the author claimed 2 bp and the market claimed 20 bp is similar to the equity premium puzzle using the CCAPM approach. Milevsky, Promislow and Young (2005, 2006) proposed a Sharpe ratio approach, which is based on mean and volatility of payments instead of returns. The methodology used in this paper is the equivalent utility pricing principle. Our approach is suitable for pricing in incomplete market. It provides a narrow price range for negotiation. The resulting risk premia are consistent with other financial market risk premia (like inflation risk premium). Our pricing framework is flexible in handling different payoff structures, basis risk and natural hedging possibilities.

Apart from securitization, there are three other possibilities of managing longevity risk,[5] namely hedging, reserving and risk sharing. The longevity risk could be partly hedged using natural hedging, for example between life annuity and term insurance. This paper illustrates the effect of the natural hedging on longevity risk premia. The impact of natural hedging is potentially significant.

The organization of this paper is the following: Section 2 introduces stochastic mortality models in order to quantify longevity risks. Section 3 describes longevity-linked securities, and incomplete-market pricing principles. In Section 4 we quantify the (seller's minimum) longevity risk premium for EIB/BNP type of longevity bonds using the equivalent utility pricing principle. Section 5 extends

[5] See also Brown and Orszag (2006), Blake, Cairns and Dowd (2006).

the longevity risk premia calculation to other longevity linked securities, including swaps, deferred bonds, floors and caps. Section 6 introduces the possibility of natural hedging into our pricing framework. Section 7 considers buyer's maximum premium, together with the presence of basis risk. Section 8 concludes.

4.2 Stochastic Mortality Models and Longevity-Linked Securities

This section first presents a stochastic mortality model to quantify longevity risks, and then describes the longevity-linked securities in more details.

4.2.1 Stochastic Mortality Models

The literature of stochastic mortality trend starts from Lee and Carter (1992).[6] According to Deaton and Paxson (2004), the Lee-Carter model has become the 'leading statistical model of mortality in the demographic literature'. Therefore, the numerical results presented in this paper are based on Lee and Carter (1992) model.

Other stochastic mortality models may also fit well in our pricing framework. Dahl (2004) and Schrager (2006) advocate the affine stochastic mortality models to capture the birth cohort mortality dynamics over one's life cycle instead of the time series of an age group over time. We leave the affine stochastic mortality approach as a future work for robustness analysis.

Lee and Carter (1992) model the time series behavior of log central death rate of an age group by using a single latent factor. The latent factor drives the mortality rates of all age cohorts. Formally, the log mortality rate of the x-year-old, $\ln\left(\mu_{x,t}\right)$, is determined by a common latent factor, γ_t, with an age-specific sensitivity parameter β_x and an age-specific level parameter α_x

$$\ln\left(\mu_{x,t}\right) = \alpha_x + \beta_x \gamma_t + \delta_{x,t} \tag{4.1}$$

[6] Various extensions of the Lee-Carter model can be found in Cairns, Blake and Dowd (2005b), Hari, De Waegenaere, Melenberg and Nijman (2006).

where transitory shocks $\delta_{x,t} \overset{iid}{\sim} N\left(0, \sigma_\delta^2\right)$ are white noises. And γ_t satisfies a random walk with drift process as

$$\gamma_t = c + \gamma_{t-1} + \varepsilon_t \tag{4.2}$$

where permanent shocks $\varepsilon_t \overset{iid}{\sim} N\left(0, \sigma_\varepsilon^2\right)$ are white noises. $\delta_{x,t}$ and ε_t are independent.

Assuming that the force of mortality is constant during a year $\mu_{x+u,t+u} = \mu_{x,t}$ $(0 \leq u < 1)$, the survival probability at time t of the x-year-old over one-year horizon is given by $p_x(t) = p_{[x_0]+t}(t) = \exp\left(-\mu_{x,t}\right)$. Similarly, the conditional probability at time t of an x-year-old surviving next τ years is given by

$$_\tau p_x(t) = \exp\left(-\sum_{i=1}^{\tau} \mu_{x+i,t+i}\right) \tag{4.3}$$

We estimate this model using the yearly U.K. (England and Wales) and Dutch male mortality data from 1880 to 2003, downloaded from the Human Mortality Database.[7] Appendix A provides a more detailed treatment of the model, together with its estimation and simulation procedures. The estimation results using the British data are included in the main text below, while the estimation results using the Dutch data are reported in Appendix A.

The estimated latent process (in the United Kingdom) is the following, including two temporary shocks captured by a 'WWI' dummy and a 'WWII' dummy[8]:

$$\gamma_t = -0.0725 + \gamma_{t-1} + 0.65 * WWI_t + 1.9 * WWII_t + \varepsilon_t \tag{4.4}$$

with the volatility $\widehat{\sigma}_\varepsilon = 0.169$. The dummy variables do not change the drift but reduce the volatility of the innovations.

[7] Human Mortality Database University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on March 27, 2006).

[8] The 'WWI' dummy takes non-zero values for years {1914 =1.5; 1915 = 1.5; 1916 = 1; 1917 = 1; 1918 = 1; 1919 = -4; 1920 = -2} and zero elsewhere. The 'WWII' dummy takes non-zero value for years {1940 = 1; 1944 = 0.1; 1945 = 0.1; 1946 = -1.2} and zero elsewhere.

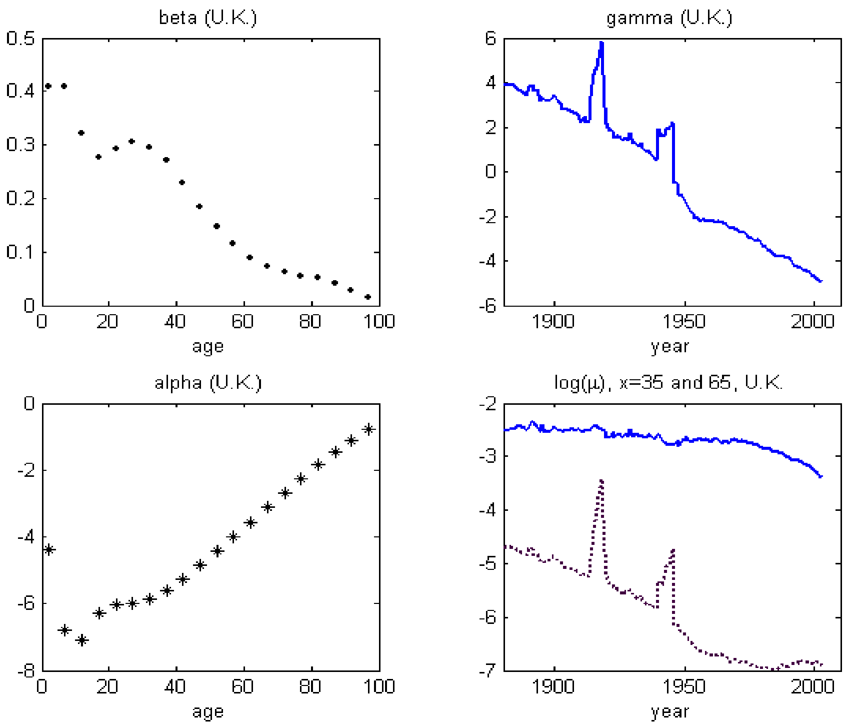


FIGURE 1: The estimated Lee-Carter (1992) model parameters, with β_x (left upper panel), γ_t (right upper panel), a_x (left lower panel) and $\ln \mu_x$ (right lower panel, the top curve for $x = 65$, and the bottom dashed curve for $x = 35$).

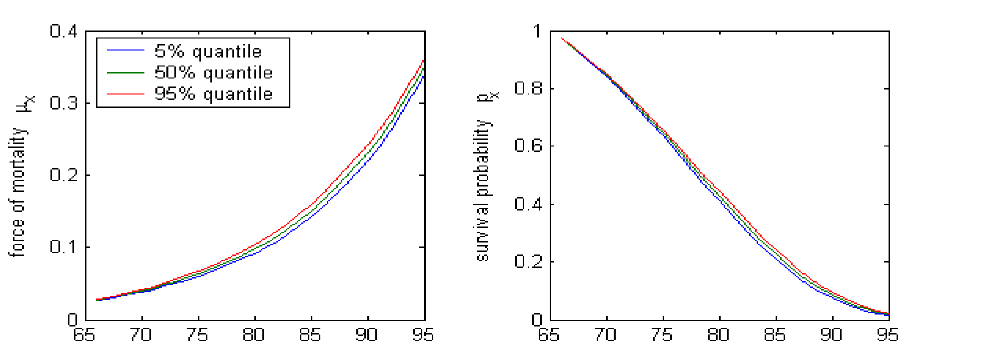


FIGURE 2 (a): The forecasted force of mortality μ_{x+t} (left panel) and the forecasted survival probabilities ${}_tp_x$ (right panel) of the 65-year-old cohort retiring in year 2004.

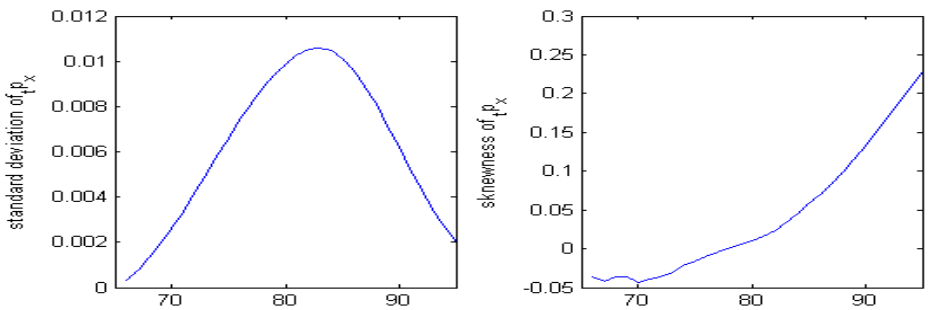


FIGURE 2 (b): The standard deviation (left) and skewness (right) of the simulated survival probability, ${}_tp_{65}$.

Assuming that the estimated model (4.4) is the 'true' process, and taking the estimated parameters as of year 2003, we could simulate the latent process γ_t and the resulting survival probabilities of the 65-yaer-old male cohort from year 2004 onwards. Figure 2(a) describes the distribution of a set of simulated survival probabilities of the 65-year-old male. From Figure 2(b) we see that the volatility of the survival probabilities exhibits a hump shape, which means that the uncertainty over a longer horizon (up to 20 years) first increases and then decreases. The distribution is also skewed. The skewness increases with age.

Using the estimation results, we can show the size of the uncertainties involved in life expectancy and annuity prices. The expected remaining life time of an x-year-old at time t is given by

$$E_t\left[T\right] = E_t\left[\sum_{\tau=1}^{\omega-x}\left({}_{\tau}p_x\left(t\right)\right)\right] \tag{4.5}$$

where ω denotes the maximum obtainable age, e.g., $\omega = 110$. The price of an immediate annuity paying 1 euro in each surviving year, assuming a fixed and flat yield curve r, is given by

$$E_t\left[L\right] = E_t\left[\sum_{\tau=1}^{110-x} e^{-rt}\left({}_{\tau}p_x\left(t\right)\right)\right]. \tag{4.6}$$

According to the estimated Lee-Carter model, the remaining lifetime of a 65-year-old British male in year 2004, is 16 years with standard error of [± 0.2] years, as given by (4.5). An immediate annuity paying 1 euro in each surviving year on average worthies[9] 13.1 euro, assuming a fixed and flat yield curve at $r = 2\%$, as given by (4.6). The standard error of the value of this annuity is [± 0.15] euro, or [± 1.1%] in relative terms. The longevity risk in these immediate annuities is not negligible.

4.2.2 Longevity-Linked Securities

Given the potential size of the longevity trend uncertainty, financial markets proposed longevity-linked securities. The first longevity bond was announced by the EIB and BNP Paribas in November 2004, but withdrawn for redesign in late 2005. The EIB/BNP bond is a ‘coupon-based’ bond, in which the notional annual coupon is indexed to a cohort survivor index in England and Wales. This cohort retires at age 65 in 2004. The maturity of this bond is 25 years. Section

[9] Money’s worth of annuity is the expected discounted value of future payments, without risk loadings.

4 discusses the pricing of such bonds. Blake, Cairns and Dowd (2006) address the lessons learned from the failure of the EIB/BNP survivor bond and provide constructive suggestions for future developments of the flourishing new market. The main lessons are the following:

1) The designed 25-year horizon is perhaps too short for an effective hedge, since longevity risk in the near future (<10 years) is small.

2) The up-front capital requirement is large, especially since a major part of the capital is taken by the ineffective hedge coupons in the near future.

3) The coupons are indexed to 65-year-old males, but annuity providers worry about longevity risk of younger cohorts and females.

4) There is large uncertainty about what the right price is that should be charged.

5) Hedge failure or basis risks are large, due to a number of ways: the reference population is different from that of an annuity provider, the survivor index is not timely available, etc.

6) The payments are nominal, whereas most pension schemes aim at inflation-linked real payments.

Their paper also introduces a few innovative hypothetical mortality-linked securities as potential solutions to the aforementioned problems. These securities include Zero-Coupon Longevity Bonds, Longevity Bull Spread Bonds, Deferred Longevity Bonds, Vanilla Mortality Swaps, Survivor Caps and Floors, Mortality Swaptions and Longevity Future. Enlightened by these discussions, this paper presents the required longevity risk premia in different longevity-linked security designs, including zeros, EIB/BNP bonds, swaps, caps and floors and deferred longevity bonds. This paper also considers the impacts of natural hedge and basis risk in Sections 6 and 7.

4.3 Longevity Risk Pricing Principles, A Review

This section reviews the recent literatures on incomplete market pricing, and motivates our choice of the methodology. Finally, we specify the utility preference needed for the equivalent utility pricing principle.

CAPM- and CCAPM-based approach. Friedberg and Webb (2005) apply Capital Asset Pricing Model (CAPM) and Consumption-based Capital Asset Pricing Model (CCAPM) to estimate the premium for longevity risk. The authors construct a pseudo-EIB/BNP survivor bond. Let $R_{b,t}$ denote the returns of such pseudo survivor bonds. Following the CAPM, the longevity risk premium is its beta, which is defined by $\beta_b = cov\left(R_b, R_m\right)/\sigma_m^2$, times the market risk premium:

$$E_t(R_b) - R_f = \beta_b\left[E\left(R_m\right) - R_f\right]$$

The authors claim that the beta on pseudo-EIB/BNP bond is 0.15 with 95 percent confidence interval of [-0.15, 0.46]. Therefore, if market risk premium is 5 percent, the longevity risk premium on this bond is 75 basis points, with confidence interval of [-75, 230] basis points. Given the wide confidence interval, the authors suggest that CCAPM as a better alternative.

Following the CCAPM, the longevity risk premium is determined by the relationship between the expected return on the asset and the marginal utility of consumption.

$$E_t(R_{b,t+1}) - R_f = -\frac{Cov_t(U'(C_{t+1}), R_{b,t+1})}{E_t(U'(C_{t+1}))}$$

The paper shows that the correlation between consumption growth and survivor bond returns is -0.1958 and is significant. However, since the standard deviation of mortality bond returns is small, as a result, the covariance between survivor bond returns and consumption growth is extremely small at -0.0015 percent. Applying the CCAPM, the risk premium is only two basis points when the coefficient of risk aversion equals 10. This result is far below the 20 bp risk premium marketed in the EIB/BNP bonds.

Sharpe ratio approach. Cochrane and Saa-Requejo (2000) suggest that the absolute value of the Sharpe ratio on any unhedgeable portfolio should be bounded, so that too 'good deals' are ruled out. Milevsky, Promislow and Young (2005, 2006) propose a so-called instantaneous Sharpe ratio to determine the mortality risk premia. Using the analogy to the Sharpe ratio in the financial market, which is the ratio of the expected excess return and the return volatility $SR^{Market} :=$

$(E[R] - R_f) / \sigma[R]$, the Sharpe ratio in the insurance context could be defined as the excess payoff above the expected payment, divided by the standard deviation of the risky payment, $SR^{Insur} := (N(1+L) - E[W_N]) / \sigma[W_N]$. The authors argue that the longevity risk loading L will be set so that the Sharpe Ratio is consistent with other asset classes in the economy. For example, if the Sharpe ratio for large cap equities is roughly 0.25, then the Sharpe ratio of the insurance policy should also be bounded within a similar magnitude.[10]

Equivalent utility based approach. The pricing method proposed in this paper is based on the equivalent utility pricing principle. The equivalent utility based approach is a popular pricing methodology for incomplete market setting. The related literature includes Svensson and Werner (1993), Young and Zariphopoulou (2002), Young (2004), Musela and Zariphopoulou (2004), De Jong (2007), Chen, Pelsser and Vellekoop (2007) and other references listed in the bibliography. As pointed out by Svensson and Werner (1993), the shadow value of a non-traded or non-hedgeable asset (the price of longevity risks in this case) can be interpreted as an additional amount of wealth added to the investor's budget so that the investor is indifferent between holding a non-hedgeable asset and hedgeable asset. Furthermore, the shadow value is investor-specific, depending on the investor's preference. In the context of longevity-linked securities, equivalent utility pricing principle reveals the minimum compensation required by the seller and the maximum price acceptable to the buyer. Therefore, this paper shows the range of possible prices for the longevity-linked securities before the market opens up.

De Jong (2007) applies the principle of equivalent utility in pricing wage-linked securities, in an incomplete market setting. In the context of defined benefit pension fund liability valuation, the main source of unhedgeable risk is the real wage growth. The pension fund is modeled as a potential buyer of the wage-linked bonds. Hence the equivalent utility pricing gives the maximum risk premium that the pension fund is willing to pay in order to obtain the insurance against the wage rate fluctuations. The paper shows the risk premium is determined by the

[10] The authors are still working on the estimation of the Sharpe ratio using annuity rate quotes. The results are not available yet.

additional wealth needed to be invested in the financial market in order to provide the participants the same level of utility as a fully wage-indexed pension.

Assuming exponential utility function, Musela and Zariphopoulou (2004), as well as Henderson (2002), show a simple analytical formula for pricing an non-traded claim. As we shall see, the pricing formula of the longevity-linked claims derived in this paper is consistent with the result found by above mentioned authors. The next sub-section illustrates the idea of equivalence pricing principle using a very simple model. The complete model is treated in Section 4.

4.3.1 An Application

Before going into the pricing, let us fix some notations. In the context of annuities, let N denote the initial size of the x-year-old cohort at time zero, and K is the agreed amount of annuity payment per annual. In the context of longevity bonds with varying coupons, NK denotes the notional coupon, and $NK_t p_x$ is the actual amount of coupon due in year t. In this paper, K is normalized to 1. Finally, the number of survivors in this cohort in year t is given by $S_t = NK_t p_x$. The survival index in t years' time is a random variable, with mean $E[S_t]$, and variance $Var[S_t]$. We assume that the longevity risk is the only risk factor in this simple illustration.

Now we illustrate the equivalent pricing principle by pricing a zero coupon longevity bond, with maturity of t years. Such zero coupon bond is effectively a large group of single premium endowment contracts which pays an agreed amount (normalized to $K = 1$) at a future time t to the survivors of the current x-year-old cohort. The longevity risk can be described as the deviation from the expected survival rate, $S_t - E[S_t]$. Let's call the longevity bond issuance company (or the pure endowment seller) the 'seller', since the 'seller' provides insurance against longevity risk. The single premium is paid at time zero, and consists of two parts. One part is the expected loss $E[S_t]$; the other part is a risk premium loading P.

The seller invests her initial wealth W_0 and the received total premium $(E\,[S_t]+P)$ in risk free asset. Further assume that the risk free rate is zero, hence, $W_t = W_0$. The minimum premium loading for this single premium endowment contract is the lowest amount that the seller asks for bearing the longevity risk $S_t - E[S_t]$. Thus, the minimum premium loading, denoted as P^-, equalizes the expected utility of underwriting the risk S_t with a compensation $E[S_t] + P^-$, with the utility of not underwriting the risk, from the seller's viewpoint. Let $U(\cdot)$ denote the utility function of the seller, we have

$$E\left[U\left(W_t + E[S_t] + P^- - S_t\right)\right] = U(W_t)$$

Case 1: CARA utility

First we assume the seller has a constant absolute risk averse (CARA) utility: $U(w) = -\frac{1}{\overline{\alpha}}\exp(-\overline{\alpha}w)$, where $\overline{\alpha}$ is the absolute risk aversion coefficient.

$$\begin{aligned}
E\left[U\left(W_t + E[S_t] + P^- - S_t\right)\right] &= -\frac{1}{\overline{\alpha}}\exp(-\overline{\alpha}W_t) \\
e^{\left(-\overline{\alpha}(E[S_t]+P^-)\right)}E[e^{(\overline{\alpha}S_t)}] &= 1 \\
P^- &= \frac{1}{\overline{\alpha}}\ln E\left[\exp\left(\overline{\alpha}\left(S_t - E[S_t]\right)\right)\right] \qquad (4.7)
\end{aligned}$$

The resulting risk loading in expression (4.7) is the so-called exponential risk premium (Kaas, et al. (2001), p. 7). The total premium can be seen as the 'best estimate' plus the (macro) risk premium which equals the logarithm of the moment generating function of risk S_t at argument $\overline{\alpha}$ divided by the CARA coefficient $\overline{\alpha}$. Notice that the risk loading is not affected by the initial wealth, W_0, for the CARA preference.

In a special case, if S_t is normally distributed, then the minimum loading P^- is proportional to the variance of S_t, as given in expression (4.8). However, Figure 2(b) shows that the distribution of S_t is skewed. Therefore the handy expression (4.8) is not an accurate approximation.

$$\begin{aligned}
\text{if} \qquad S_t &\sim N\left(E[S_t],\ Var[S_t]\right), \\
\text{then} \qquad \overline{\alpha}\left(S_t - E[S_t]\right) &\sim N\left(0,\ \overline{\alpha}^2 Var[S_t]\right) \\
E\left[\exp\left(\overline{\alpha}\left(S_t - E[S_t]\right)\right)\right] &= \exp\left(\frac{1}{2}\overline{\alpha}^2 Var[S_t]\right) \\
P^- &= \frac{1}{2}\overline{\alpha}Var[S_t] \qquad (4.8)
\end{aligned}$$

Case 2: CRRA utility

Alternatively we assume the seller has a constant relative risk averse (CRRA) utility: $U(w) = w^{1-\gamma}/(1-\gamma)$. Notice that the risk loading does depend on the initial wealth and risk aversion parameter for the CRRA preference. The minimum loading P^- is the one that solves Equation (4.9).

$$\begin{aligned} E\left[U\left(W_t + E\left[S_t\right] + P^- - S_t\right)\right] &= \frac{W_t^{1-\gamma}}{1-\gamma} \\ E\left[\left(W_t + E\left[S_t\right] + P^- - S_t\right)^{1-\gamma}\right] &= W_t^{1-\gamma} \\ E\left[\left(1 + \frac{E\left[S_t\right] + P^- - S_t}{W_t}\right)^{1-\gamma}\right] &= 1 \end{aligned} \tag{4.9}$$

Numerical results:

The risk loading P^- of both cases can be evaluated by means of simulation. Using the estimation and simulation procedures presented in Appendix A, we calculate the expected loss $E\left[S_t\right]$ and the risk loading P^- for the endowment contract. The total premium paid per person is $\frac{E[S_t]+P^-}{N}$. We can express P^- in terms of risk premium, R_p, which is a discount rate above the risk free rate (i.e. 0 percent as assumed) and an actuarial discount rate R_a:

$$\frac{1}{\left(1 + R_a + R_p\right)^t} = \frac{E\left[S_t\right] + P^-}{N}$$

where the actuarial discount rate R_a is defined by

$$\frac{1}{\left(1 + R_a\right)^t} = \frac{E\left[S_t\right]}{N} = E[{}_tp_x]$$

The risk premia for CARA and CRRA utility specifications are presented in Tables 1 and 2 below.

	N = 10			N = 100			N = 1000		
maturity	CARA = 1	CARA = 3	CARA = 5	CARA = 1	CARA = 3	CARA = 5	CARA = 1	CARA = 3	CARA = 5
5	0	0	0	-1	-2	-3	-5	-12	-15
10	0	-1	-1	-2	-7	-11	-18	-24	-26
15	-1	-2	-3	-6	-16	-24	-31	-38	-39
20	-1	-3	-5	-10	-28	-39	-48	-55	-57
25	-1	-4	-6	-13	-37	-54	-70	-82	-85
30	-1	-3	-4	-9	-27	-46	-78	-104	-109
35	0	-1	-1	-3	-9	-15	-31	-91	-112

TABLE 1: The longevity risk premium R_p in basis points, for different cohort sizes (N=10, 100, 1000 $\times 10^3$) and different risk aversion values ($CARA(\alpha) = 1, 3, 5$).

	N = 10			N = 100			N = 1000		
maturity	CRRA = 3	CRRA = 5	CRRA = 8	CRRA = 3	CRRA = 5	CRRA = 8	CRRA = 3	CRRA = 5	CRRA = 8
5	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	-1	-1	-2
15	0	0	0	0	0	0	-2	-3	-5
20	0	0	0	0	-1	-1	-3	-5	-8
25	0	0	0	0	-1	-1	-4	-6	-10
30	0	0	0	0	0	-1	-3	-4	-7
35	0	0	0	0	0	0	-1	-1	-2

TABLE 2: The longevity risk premium R_p in basis points, for different cohort sizes (N=10, 100, 1000 $\times 10^3$) and different risk aversion values ($CRRA(\gamma) = 3, 5, 8$). The initial wealth of the insurer is assumed to be $W_0 = 100$.

Key features of the results are: 1) The required risk premium is negative, meaning that the bond yield is lower than the risk free rate, so that the bond price is higher than the risk free bond. Thus, the insurer (i.e., the survivor bond issuer) is compensated for bearing longevity risks. 2) The additional discount rate R_p (in absolute value) increases as the size of the pool increases. 3) The more risk averse the insurer is, the higher compensation is required. 4) Approximately, the longer the maturity, the higher compensation is required.

4.3.2 Preference Assumption

The results in Tables 1 and 2 reveal some unsatisfied properties of CARA and CRRA preferences. For a CARA investor, he has the same worry about one additional euro loss no matter how rich or poor he is. For a CRRA investor, he cares much less when he is rich. In our view, both preferences are too restrictive to characterize the risk preference of financial institutions. The CARA utility might overestimate the longevity risk premium. Whereas, the CRRA utility might underestimate the risk premium. Therefore, we modify the CARA utility, to make the risk aversion depend on the initial capital or initial wealth of the company, $\alpha(W_0) = \overline{\alpha} W_0^{-b}$, where $b \in [0,1]$. In general, risk aversion decreases with initial wealth.[11] At the two extremes, the modified utility approaches CRRA specification when $b = 1$, and the modified utility is back to the CARA specification when $b = 0$. The proposed preference (4.10) retains some convenient features of the negative exponential utility function, since it is separable for independent risks x and y, as $E[u(x+y)] = E[u(x)]\, E[\exp(-\alpha y)]$.

$$u(S) = -\frac{1}{\alpha(W_0)} \exp(-\alpha(W_0) S) \tag{4.10}$$

It is important to be clear about whose preference that (4.10) captures. Sections 3 to 5 focus on seller's minimum required risk premium. Therefore the utility function (4.10) represents the preference of the shareholders of the seller. In the context of the EIB/BNP longevity bond, it is the preference of the shareholders of EIB/BNP. Section 6 discusses the buyer's maximum risk premium. Hence the utility function (4.10) represents the preference of the buyer, e.g., a pension fund.

4.4 Pricing of a Coupon-Based Longevity Bond

This section derives the minimum required longevity risk premium of a coupon based longevity bond from the seller's point of view. The longevity bond is linked to a large pool of population. Therefore mortality risk (also called micro longevity risk) is fully diversified. The setup of the model is the following. The shareholder

[11] Wachter and Jogo (2007) (and their references) provide arguments and evidences for a wealth-varying risk aversion.

of a financial company (like EIB/BNP) derives her utility from dividends and final wealth. We consider two alternative situations. In the first situation, the company does not insure longevity risk, hence is not exposed to longevity risk. In the second situation, the company issues a longevity bond and hence bears longevity risk. Furthermore, we assume that the macro longevity risk is independent from the financial risks. The methodology used in this section combines the martingale approach with the equivalent utility pricing principle.

4.4.1 Setup

Problem 1 *without longevity risk*

Assume a complete financial market, with constant risk free rate, r. The shareholder (or manager) of the company derives her utility from dividends and final wealth at the end of the horizon, T. The per period utility is described as (4.10), and δ is the subjective discount rate of the shareholders. The initial equity capital of the company is given by W_0. The company maximizes the shareholder's utility by optimizing asset allocation (x_t) and dividend (D_t) decisions. Formally the optimization problem is the following

$$\max_{\{D_t,x_t\}_{t=0}^{T},W_T} V_0 = E\left[\int_0^T e^{-\delta t}u\left(D_t\right)dt + e^{-\delta T}u\left(W_T\right)\right] \tag{4.11}$$

$$s.t. \qquad E\left[\int_0^T M_t D_t dt + M_T W_T\right] = W_0 \tag{4.12}$$

where M_t is the pricing kernel for the complete financial market. M_t is defined by

$$dM_t/M_t = -rdt - \lambda dZ_t \tag{4.13}$$

where λ is the market price of equity risk.

Problem 2 *with longevity risk*

In the same financial market, the company issues coupon-based longevity bond, in which the annual coupon is indexed to the 1939-born cohort survivor index. This cohort retires in 2004 at age 65. The longevity risk is not hedgeable from the financial market. Hence, the company derives her utility from dividends and the residual claim $(E[S_t]-S_t)$ from the longevity risk. The initial equity capital of the company is now augmented by an additional risk loading π. The company maximizes her utility by optimizing asset allocation and dividend decisions. Formally the optimization problem is the following

$$\max_{\{D_t^\pi, x_t\}_{t=0}^T, W_T^\pi} V_0^\pi = E\left[\int_0^T e^{-\delta t} u\left(D_t^\pi + E[S_t] - S_t\right) dt + e^{-\delta T} u\left(W_T^\pi\right)\right] \tag{4.14}$$

$$s.t. \qquad E\left[\int_0^T M_t D_t^\pi dt + M_T W_T^\pi\right] = W_0 + \pi \tag{4.15}$$

Applying the equivalent utility pricing argument, we determine the minimum risk compensation π such that the company is indifferent from bearing the longevity risk and without the longevity risk. That is, the indirect utility must be equal under these two situations:

$$V_0^\pi(\pi) = V_0. \tag{4.16}$$

4.4.2 Results

The derivations are given in Appendix B. The main results are the following: The risk loading π is a present value of the certainty equivalent compensations for the risks $S_t - E[S_t]$, as given by (4.17).

$$\begin{aligned} \pi &= \frac{1}{\alpha}\int_0^T e^{-rt} \ln E\left[\exp\left(-\alpha\left(E[S_t]-S_t\right)\right)\right] dt && (4.17)\\ &= \frac{1}{\alpha}\int_0^T e^{-rt} \ln G_t dt && (4.18) \end{aligned}$$

where G_t denotes

$$G_t \equiv E\left[\exp\left(-\alpha\left(E[S_t]-S_t\right)\right)\right]$$

The value of the coupon-based longevity bonds with maturity T can be decomposed into 'best estimated' price ($\int_0^T e^{-rt}E[S_t]dt$) and longevity risk loading (π):

$$total\ price = best\ estimate\ +\ risk\ loading = \int_0^T e^{-rt}E[S_t]dt + \pi \tag{4.19}$$

Table 3 shows the normalized risk loading, $\frac{\pi}{beste\ stimate}$, of the coupon-based longevity bonds with maturity $T = 5, ..., 35$ years. The risk loading depends on the maturity of the bond, the size of initial equity capital and the risk aversion of the insurer. Recall that the risk aversion is inversely related to the size of the capital, as $\alpha(W_0) = \overline{\alpha}(W_0)^{-b}$. When b changes from 1 to 0, the preference shifts from CRRA to CARA, which results in an increase in risk loading. CRRA investor requires virtually zero risk compensation. Whereas CARA investor requires a sizable compensation, up to 1.6 percent of the best estimated cost.

We can also express the risk loading π in terms of risk premium, R_p, which can be seen as an additional discount rate above the risk free rate to adjust for the longevity risk. As explained in Section 3.1, the risk premium is negative, meaning that the bond yield is lower than the risk free rate, so that the bond price is higher than the risk free bond. Thus, the insurer (i.e., the survivor bond issuer) is compensated for bearing longevity risks.

$$\begin{aligned} total\ price &= \int_0^T e^{-(r+R_p)t}E[S_t]dt \\ i.e. \quad \int_0^T e^{-rt}E[{}_tp_{65}]dt + \frac{\pi}{N} &= \int_0^T e^{-(r+R_p)t}E[{}_tp_{65}]dt \end{aligned}$$

equity	w0 = 10000				w0 = 1000				w0 = 100			
maturity	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0
5	0%	0%	0,0%	0,0%	0%	0,0%	0,0%	0,0%	0%	0,0%	0,0%	0,0%
10	0%	0%	0,1%	0,2%	0%	0,0%	0,1%	0,2%	0%	0,1%	0,1%	0,2%
15	0%	0%	0,2%	0,5%	0%	0,1%	0,2%	0,5%	0%	0,2%	0,3%	0,5%
20	0%	0%	0,3%	0,9%	0%	0,2%	0,4%	0,9%	0%	0,3%	0,6%	0,9%
25	0%	0%	0,4%	1,4%	0%	0,3%	0,6%	1,4%	0%	0,4%	0,8%	1,4%
30	0%	0%	0,5%	1,5%	0%	0,3%	0,7%	1,5%	0%	0,5%	0,9%	1,5%
35	0%	0%	0,5%	1,6%	0%	0,3%	0,7%	1,6%	0%	0,5%	0,9%	1,6%

TABLE 3: The normalized risk loading per person, $\pi/best\ estimate$, of EIB/BNP longevity bonds for different sizes of the initial equity capital, $W_0 = [10000, 1000, 100]$ million, and different risk aversion specifications $\alpha(W_0) = \overline{\alpha}(W_0)^{-b}$, with $\overline{\alpha} = 3$, $b = [1, \frac{1}{4}, \frac{1}{8}, 0]$. The size of the insured pool is $N = 100$ million, $K = 1$.

equity	w0 = 10000				w0 = 1000				w0 = 100			
maturity	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0
5	0	0	0	-1	0	0	0	-1	0	0	-1	-1
10	0	0	-1	-3	0	-1	-1	-3	0	-1	-2	-3
15	0	-1	-2	-7	0	-1	-3	-7	0	-2	-4	-7
20	0	-1	-4	-11	0	-2	-5	-11	0	-4	-7	-11
25	0	-2	-5	-15	0	-3	-6	-15	0	-5	-8	-15
30	0	-2	-5	-16	0	-3	-7	-16	0	-5	-9	-16
35	0	-2	-5	-16	0	-3	-7	-16	0	-5	-9	-16

TABLE 4: The longevity risk premium R_p (in basis points) of EIB/BNP longevity bonds for different sizes of the initial equity capital, $W_0 = [10000, 1000, 100]$ million, and different risk aversion specifications $\alpha(W_0) = \overline{\alpha}(W_0)^{-b}$, with $\overline{\alpha} = 3$, $b = [1, \frac{1}{4}, \frac{1}{8}, 0]$. The size of the insured pool is $N = 100$ million, $K = 1$.

Table 4 presents the risk premium R_p (in basis points) of the 'coupon-based' longevity bonds with maturity $T = 1, ..., 35$ year. The results show two things. First, the risk premium increases as the maturity of the bond increases. The risk premium for short maturity ($T \leq 5$ years) is small, less than one basis point. The minimum risk premium for long maturity ($T = 30$) is around 7 to 9 basis points (taking $b = 1/8$). Second, the risk premium depends on the financial position of the insurer. The larger initial equity a firm has, the lower risk premium the firm requires (except for the CARA case ($b = 0$)). The face value of the EIB/BNP bond issue was 540 million and the bond had a 25-year maturity. The initial coupon was set at 50 million, which is comparable with the initial payments $NK = 100$ million assumed here. By the end of 2005, EIB's own fund amounts to nearly 30000 million, which is comparable with the initial equity $W_0 = 10000$ million assumed here. The left panel with $W_0 = 10000$ indicates a (sell-side) minimum required risk premium of five basis points with $b = 1/8$ for 25 years maturity.

4.4.3 Implications

The implication that we can get from the above results is that longevity risk premium depends on the financial position of the insurer. Large equity financial institutions may require a lower risk premium. Put differently, smaller issues (smaller K) may require lower risk premium. In order to avoid too high risk premia, it might be helpful to have many large institutions all issuing moderate amounts of longevity bonds, linked to the same survivor index.

4.5 Pricing of Other Longevity-Linked Securities

In this section, we look at other types of longevity-linked securities, including swaps, deferred starting bonds, floors and caps. Since the market is incomplete, we will show that different payoff structures may lead to different risk premia.

4.5.1 Vanilla Longevity Swaps

Vanilla longevity swaps have the same risk structure, $E[S_t] - S_t$, as the longevity bonds. The insurer or the investment bank pays the counterpart the difference between the expected and realized mortality. Analogize to interest rate swap, the fixed leg is $E[S_t]$, and the floating leg is S_t. The required risk premium of a longevity swap is the same as in a longevity bond with the same maturity and the same amount of notional issues (Table 4). The main advantages of swap lie in a much lower up-front capital requirement and lower credit risk, as compared with a long maturity longevity bond.

4.5.2 Deferred Longevity Bonds

A deferred longevity bond starts paying the longevity-linked coupons s years after the issuance, till the bond maturity in year T. An advantage of a deferred longevity bond is that it skips the ineffective hedge coupons in the first few years, and hence requires much less up-front capital than an immediate coupon paying bond. Following the same pricing principle as in Section 4, the risk loading of a deferred longevity bond is given by

$$\pi^{def} = \frac{1}{\alpha} \int_s^T e^{-rt} \ln G_t dt \tag{4.20}$$

where α is the shorthand notation for $\alpha\left(W_0\right) = \overline{\alpha} W_0^{-b}$ and $G_t \equiv E\left[\exp\left(-\alpha\left(E[S_t] - S_t\right)\right)\right].$

# year defer	best estim.	w0 = 10000				w0 = 1000				w0 = 100			
		b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0
0	13,1	0%	0%	1%	2%	0%	0%	1%	2%	0%	1%	1%	2%
5	8,6	0%	0%	1%	2%	0%	0%	1%	2%	0%	1%	1%	2%
10	5,2	0%	0%	1%	4%	0%	1%	2%	4%	0%	1%	2%	4%
15	2,7	0%	1%	2%	6%	0%	1%	2%	6%	0%	2%	3%	6%
20	1,2	0%	1%	2%	8%	0%	1%	3%	8%	0%	2%	4%	8%

TABLE 5: The normalized risk loading, $\pi/best\ estimate$, of the deferred longevity bonds, for different sizes of the initial equity capital, $W_0 = [10000, 1000, 100]$ million, and different risk aversion values $\overline{\alpha}\,(W_0)^{-b}$, with $\overline{\alpha} = 3,\ \ b = [1, \frac{1}{4}, \frac{1}{8}, 0]$. $N = 100$ million, $K = 1$.

# year defer	w0 = 10000				w0 = 1000				w0 = 100			
	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0
0	0	-2	-5	-16	0	-3	-7	-16	0	-5	-9	-16
5	0	-2	-6	-17	0	-3	-8	-17	0	-6	-10	-17
10	0	-2	-7	-21	0	-4	-9	-21	0	-7	-12	-21
15	0	-3	-9	-26	0	-5	-12	-26	0	-9	-15	-26
20	0	-3	-10	-30	0	-6	-13	-30	0	-10	-18	-30

TABLE 6: The longevity risk premium R_p (in basis points) of deferred longevity bonds, for different sizes of the initial equity capital, $W_0 = [10000, 1000, 100]$ million, and different risk aversion values $\overline{\alpha}\,(W_0)^{-b}$, with $\overline{\alpha} = 3$, $b = [1, \frac{1}{4}, \frac{1}{8}, 0]$. $N = 100$ million, $K = 1$.

Tables 5 and 6 show the relative risk loadings and the risk premia of several deferred longevity bonds. The following results assume that all deferred longevity bonds mature in 35 years, but the first coupon payments could start 5, 10, 15 or 20 years after the issuance. Notice the first row in the tables is an immediate starting bond for comparison. The initial capital is much smaller than the immediate starting bond, but the relative risk loading is much larger. As a consequence, the required risk premia are also larger than the immediate starting bond.

4.5.3 Longevity Floors and Longevity Caps

For the case of a longevity floor, the payoff structure is $\min[E[S_t] - S_t, 0]$. When the number of survivor is greater than the expected, the insurer faces a 'longevity' loss. The payoff structure of a longevity cap is $max[E[S_t] - S_t, 0]$. The insurer makes 'longevity' profit when the number of survivor is less than the expected. The risk premium of longevity floor and cap can be obtained in similar way based on the equivalent utility approach. The company derives her utility from dividends and the residual claim $([E[S_t] - S_t]^-)$[12]. Following the martingale approach, the risk loading is determined by

$$\pi^- = \frac{1}{\alpha}\int_0^T e^{-rt}\ln G_t^- dt \tag{4.21}$$

$$\pi^+ = \frac{1}{\alpha}\int_0^T e^{-rt}\ln G_t^+ dt \tag{4.22}$$

where $G_t^- \equiv E\left[\exp\left(-\alpha\left[E[S_t] - S_t\right]^-\right)\right]$, and $G_t^+ \equiv E\left[\exp\left(-\alpha\left[E[S_t] - S_t\right]^+\right)\right]$.

Table 7 compares the longevity risk premium R_p (in basis points) of longevity bonds, longevity floors and longevity caps respectively.

	w0 = 10000			w0 = 1000			w0 = 100		
maturity	bond	floor	cap	bond	floor	cap	bond	floor	cap
5	0	-3	3	0	-3	3	-1	-3	3
10	-1	-5	4	-1	-5	4	-2	-5	4
15	-2	-7	5	-3	-7	5	-4	-8	5
20	-4	-9	7	-5	-10	6	-7	-11	6
25	-5	-11	8	-6	-12	7	-8	-14	7
30	-5	-12	8	-7	-13	8	-9	-15	8
35	-5	-12	9	-7	-13	8	-9	-15	8

TABLE 7: The longevity risk premium R_p (in basis points) of longevity bonds, longevity floors and longevity caps, for different sizes of the initial equity capital, $W_0 = [10000, 1000, 100]$ million, and different risk aversion values $\overline{\alpha}\,(W_0)^{-b}$, with $\overline{\alpha} = 3$, $b = \frac{1}{8}$. $N = 100$ million, $K = 1$.

[12] $[E[S_t] - S_t]^- \equiv \min\left(E[S_t] - S_t, 0\right)$; $[E[S_t] - S_t]^+ \equiv max[E[S_t] - S_t, 0]$

We observe three things from Table 7. First, the risk premium of the longevity floor is larger (in absolute terms) than that of the longevity bond. This is because the payoff of the longevity bond is symmetric, whereas the payoff of a longevity floor is highly asymmetric. Therefore a higher risk premium is required for bearing losses only. Second, the risk premium of the long position in this longevity caps is positive, which means that the insurer pays for the call option. The more risk averse the insurer is, the less willingness to pay (read: compensate) the counterpart, for receiving the uncertain profit. Thirdly, as the initial wealth decreases (hence the relative risk aversion increases), the value of the floor increases.

4.6 The Effect of Natural Hedging

It is known that term insurance policies provide a natural hedge for the immediate annuities (see, e.g., Cox and Lin (2004)). The term insurance pays out a certain amount of death benefit if the policy holder dies before the contract expires. Since longevity shocks affect all age cohorts in the same direction, the unexpected increase in annuity payments to the retirees can be partially offset by the unexpected reduction of death benefit payments linked to the younger cohorts. The availability of natural hedging clearly affects the risk premium of the longevity bond issuance company. This section examines the magnitude of this effect on the pricing of longevity bonds.

Suppose that the longevity bond issuance company bears the risks from both the term insurance policies linked to a group of 35-year-olds in 2004 and the longevity bonds linked to a group of 65-year-olds in 2004. Further, suppose that the estimated Lee-Carter 1992 model is the true process governing the future mortality dynamics. The number of deaths for the 35-year-old cohort in year t is $S_{t-1}^{35} - S_t^{35}$. The unexpected shocks from the term insurance policies are $B_t * \left(E\left[S_{t-1}^{35} - S_t^{35}\right] - \left(S_{t-1}^{35} - S_t^{35}\right)\right)$, where B_t denotes the ratio of death benefit relative to the annuity payout K (=1, which is the agreed annuity payments). The (unexpected) shocks from the longevity bonds are captured by $E[S_t^{65}] - S_t^{65}$.The combined unexpected shocks from the term insurance and the longevity bonds can be denoted by Z_t as

$$Z_t \equiv E[S_t^{65}] - S_t^{65} + B_t * \left(E\left[S_{t-1}^{35} - S_t^{35}\right] - \left(S_{t-1}^{35} - S_t^{35}\right)\right).$$

As explained in Section 4, the minimum longevity risk loading π^{hedge} required by the seller is determined by setting $V_0^{\pi,hedge} = V_0$, where $V_0^{\pi,hedge}$ is the indirect utility given by

$$\max_{\{D_t^\pi, x_t\}_{t=0}^T, W_T^\pi} V_0^{\pi,hedge} = E\left[\int_0^T e^{-\delta t} u\left(D_t^{\pi,hedge} + Z_t\right) dt + e^{-\delta T} u\left(W_T^\pi\right)\right] \quad (4.23)$$

$$s.t. \qquad E\left[\int_0^T M_t D_t^{\pi,hedge} dt + M_T W_T^\pi\right] = W_0 + \pi^{hedge} \quad (4.24)$$

Following a similar argument as in Section 4.2, equalizing $V_0^{\pi,hedge} = V_0$, we can find the corresponding risk premium as

$$\pi^{hedge} = \frac{1}{\alpha} \int_0^T e^{-rt} \ln G_t^{hedge} dt \quad (4.25)$$

where $G_t^{hedge} \equiv E\left[\exp\left(-\alpha\left(Z_t\right)\right)\right]$, measuring the certainty equivalent of the combined shocks Z_t.

The following example illustrates the effectiveness of natural hedging. In this example, the level of death benefits linearly decreases over time,[13] i.e., $B_t = T + 1 - t$, for $t = 1, 2, ..., T$. Figure 3 compares the term insurance with the payout volatility of the longevity bonds with a term insurance with decreasing death benefits. The volatility of the combined shocks is much lower than that of the longevity bond alone. However, the hedging is not perfect. Table 8 shows the minimum required risk premia, R_p, which is clearly reduced when natural hedging is available. The risk premia are more than halved compared with the case without natural hedging.

[13] Decreasing death benefit is very common in life insurance policies bundled with mortgages for young households. As mortgages are paid off over time, the amount of death benefit decreases.

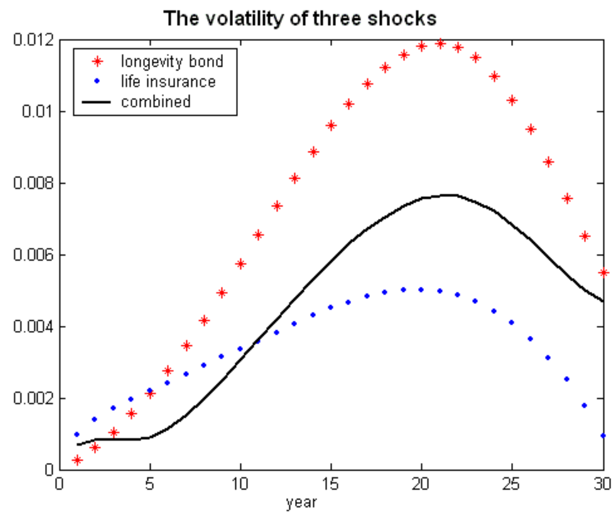


FIGURE 3: The volatility of payouts of the longevity bond and the term insurance separately and combined.

equity	w0 = 10000				w0 = 1000				w0 = 100			
maturity	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0
5	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	-1	0	0	0	-1	0	0	0	-1
15	0	0	-1	-2	0	0	-1	-2	0	-1	-1	-2
20	0	0	-1	-5	0	-1	-2	-5	0	-1	-3	-5
25	0	-1	-2	-6	0	-1	-3	-6	0	-2	-3	-6
30	0	-1	-2	-7	0	-1	-3	-7	0	-2	-4	-7

TABLE 8: The longevity risk premium R_p (in basis points) of longevity bonds when natural hedgeing is available ($B_t = T + 1 - t$).

4.7 The Demand Side Pricing and Basis Risk

4.7.1 Demand Side Pricing

The demand side pricing considers the maximum price π^{BUY} that the buyers (e.g., annuity providers) are willing to pay for the longevity bond or other securities in order to be fully insured against the longevity risk. From buyer's point of view, π^{BUY} can be derived in the same framework as in Section 4. Assume an annuity provider sold annuities to a cohort retiring in 2004 at age 65. The shareholder of this annuity provider derives her utility from dividends and final wealth. We still consider two situations. In the first scenario, the annuity provider bought the ideal EIB/BNP survivor bonds at price π^{BUY} so that the longevity risk from her annuity contracts is completely insured. In the second scenario, the annuity provider bears the longevity risk herself.

Problem 3 *without longevity risk*

Assume a complete financial market, with constant risk free rate, r. The annuity provider derives her utility from dividends and final wealth at the end of the horizon. The company bought the ideal EIB/BNP survivor bonds for π^{BUY}, such that the longevity risk is completely hedged. The company maximize her utility by optimizing asset allocation (x_t) and dividend (D_t) decisions.

$$\max_{\{D_t^\pi, x_t\}_{t=0}^T, W_T^\pi} V_0^\pi = E\left[\int_0^T e^{-\delta t} u\left(D_t^\pi\right) dt + e^{-\delta T} u\left(W_T^\pi\right)\right] \tag{4.26}$$

$$s.t. \qquad E\left[\int_0^T M_t D_t^\pi dt + M_T W_T^\pi\right] = W_0 - \pi^{BUY} \tag{4.27}$$

Problem 4 *with longevity risk*

In the same financial market, this annuity provider did not buy any longevity bond, and hence bears the longevity risk herself. The company derives her utility from dividends and the residual claim ($E[S_t] - S_t$) from the longevity risk. The longevity risk is not hedgeable from the financial market. The company maximizes her utility by optimizing asset allocation and dividend decisions.

$$\max_{\{D_t,x_t\}_{t=0}^T,W_T} V_0 = E\left[\int_0^T e^{-\delta t}u\left(D_t + E[S_t] - S_t\right)dt + e^{-\delta T}u\left(W_T\right)\right] \tag{4.28}$$

$$s.t. \qquad E\left[\int_0^T M_t D_t dt + M_T W_T\right] = W_0 \tag{4.29}$$

Applying the equivalent utility pricing argument, we want to find the minimum risk compensation π^{BUY} such that the company is indifferent from bearing the longevity risk and without the longevity risk, that is,

$$V_0^\pi(\pi^{BUY}) = V_0. \tag{4.30}$$

Following the equivalent utility pricing argument, we have

$$\pi^{BUY} = \frac{1}{\alpha}\int_0^T e^{-rt}\ln G_t^{BUY}dt. \tag{4.31}$$

where G_t^{BUY} denotes

$$G_t^{BUY} \equiv E\left[\exp\left(-\alpha\left(E[S_t] - S_t\right)\right)\right] \tag{4.32}$$

The maximum premium that a buyer of the longevity bond is willing to pay has the same form as the minimum premium that the bond issuance company requires. It is common to assume that the longevity bond buyer is more risk averse than the bond issuance company, or the financial position of the buyer is weaker than the seller.

The buyer's maximum price is also influenced by whether or not natural hedging is possible. The availability of natural hedging could reduce the buyer's price significantly. Furthermore, the presence of basis risk and the risk sharing possibility will also affect the buyer's maximum price.

4.7.2 Basis Risk

An ideal longevity bond which provides a perfect longevity hedge should be linked to the annuitant population of the annuity provider. However, quite often this is not the case. There is a discrepancy between the reference population that the bond is linked to and the annuitant population of the bond buyer. Although the survival probabilities of the two populations might be (highly) correlated, the longevity bond buyer still exposes to the remaining unhedgeable part, the so-called basis risk. As a real life example, the EIB/BNP longevity bond, although linked to the British survivor index, was also marketed among Dutch pension funds. The idea is that the Dutch survivor index may be highly correlated with the British one. The question here is whether 20 basis points is a good deal or not for Dutch pension funds. This depends on the correlation between Dutch and British mortality rates. The correlation between the innovations of the latent factors ($\triangle\gamma_t^{UK}$ and $\triangle\gamma_t^{NL}$) is about 0.8, based on 1880-2003 data from both countries, with $\triangle\gamma_t \equiv \gamma_t - \gamma_{t-1}$. The remaining part of the section examines the impact of basis risk on the pricing of longevity risk.

The basis risk between the British and the Dutch annuitant population can be captured by $Z_t^{BasisRisk}$ defined as

$$Z_t^{BasisRisk} \equiv E[S_t^{NL}] - S_t^{NL} - \left(E[S_t^{UK}] - S_t^{UK}\right).$$

Based on the expression for buyer's maximum acceptable price (4.31), we can show that the risk loading with basis risk is

$$\pi^{BasisRisk} = \frac{1}{\alpha}\int_0^T e^{-rt}\ln G_t^{BasisRisk}dt. \tag{4.33}$$

where $G_t^{BasisRisk}$ denotes

$$G_t^{BasisRisk} \equiv E\left[\exp\left(-\alpha\left(Z_t^{BasisRisk}\right)\right)\right] \tag{4.34}$$

Assume that the Dutch pension fund has the same preference and the same level of equity as the longevity bond issuance company, and also assume that the Dutch pension fund has no natural hedging possibility. If without basis risk, that is, if there were a longevity bond linked to the Dutch population, then we get the following maximum acceptable longevity risk premium R_p as given in Table 9[14].

[14]The risk premium for Dutch population (Table 9) is higher than the minimum required risk premium for British population (Table 4), due to the fact that the estimated volatility

However, since there is no such longevity bond linked to the Dutch population directly, but linked to the British population, the hedging will not be perfect. The basis risk between the two population will reduce the risk premium. Indeed, the demand side risk premium in Table 10 (with basis risk) is lower than that of in Table 9 (without basis risk).

equity	w0 = 10000				w0 = 1000				w0 = 100			
maturity	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0
5	0	0	-1	-2	0	0	-1	-2	0	-1	-1	-2
10	0	-1	-2	-7	0	-1	-3	-7	0	-2	-4	-7
15	0	-2	-5	-15	0	-3	-7	-15	0	-5	-9	-15
20	0	-3	-8	-23	0	-5	-11	-23	0	-8	-14	-23
25	0	-3	-10	-28	0	-6	-13	-28	0	-10	-17	-28
30	0	-3	-10	-30	0	-6	-14	-30	0	-10	-18	-30

TABLE 9: Without basis risk, the buyer's maximum longevity risk premium R_p (in basis points) for Dutch pension fund with different initial equity capital levels, and different risk aversion values $\overline{\alpha}\,(W_0)^{-b}$, with $\overline{\alpha} = 3$, $b = [1, \frac{1}{4}, \frac{1}{8}, 0]$. The size of the insured pool is $N = 100$ million, $K = 1$.

equity	w0 = 10000				w0 = 1000				w0 = 100			
maturity	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0	b=1	b=1/4	b=1/8	b=0
5	0	0	0	-1	0	0	0	-1	0	0	0	-1
10	0	0	-1	-2	0	0	-1	-2	0	-1	-1	-2
15	0	-1	-2	-6	0	-1	-3	-6	0	-2	-4	-6
20	0	-1	-4	-11	0	-2	-5	-11	0	-4	-7	-11
25	0	-2	-6	-16	0	-3	-8	-16	0	-6	-10	-16
30	0	-2	-7	-18	0	-4	-9	-18	0	-7	-12	-18

TABLE 10: With basis risk, the buyer's maximum longevity risk premium R_p (in basis points) for the same Dutch pension fund as in Table 9.

$\widehat{\sigma^{NL}}_{\varepsilon} = 0.2176$ is higher than the British counterpart (see Appendix A.1).

4.8 Conclusion

Longevity risk imposes serious solvency issues on pension plans and insurance companies. Longevity-linked securities are desirable instruments for buyers and sellers, but are not traded yet in financial markets because of the pricing difficulty. To tackle the pricing problem, we propose a new pricing method, which is suitable for incomplete market pricing, flexible and consistent with other financial risk premia. Our methodology is based on the equivalent utility pricing principle. The obtained range of the longevity risk premia captures the seller's minimum price and the buyer's maximum price. We apply the method in pricing various longevity-linked securities (bonds, swaps, caps and floors) linked to the U.K. and the Dutch mortality data. We show that the size of the risk premium depends on the payoff structure of the security due to the market incompleteness. Given a plausible range of risk aversion, financial position and other assumptions, we show that the resulting risk premia are consistent with the limited market observation and consistent with other financial risk premia (e.g., inflation risk premium). We also show that the impact of natural hedging is potentially significant. The results provide design implications for longevity-linked securities and longevity risk management.

4.9 Appendix A: The Lee-Carter 1992 Model

This appendix provides a more detailed treatment of the model, together with the estimation and simulation procedures. Following the Lee Carter 1992 model, the time series property of the log mortality rate of the x-year-old, $\ln\left(\mu_{x,t}\right)$, is determined by a common latent factor γ_t with an age specific sensitivity β_x and an age specific level α_x

$$\ln\left(\mu_{x,t}\right) = \alpha_x + \beta_x \gamma_t + \delta_t \tag{4.35}$$

with the latent factor satisfies a random walk with drift process as

$$\gamma_t = c + \gamma_{t-1} + \varepsilon_t \tag{4.36}$$

where δ_t and ε_t are vectors of white noise, satisfying the distributional assumptions

$$\begin{pmatrix} \delta_t \\ \varepsilon_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_\delta & 0 \\ 0 & \sigma_\varepsilon^2 \end{pmatrix}\right) \tag{4.37}$$

The forecasted log mortality rate in s years' time of a then x-year-old is

$$\begin{aligned} \ln\left(\mu_{x,t+s}\right) &= \alpha_x + \beta_x \gamma_{t+s} + \delta_{t+s} \\ &= \ln\left(\mu_{x,t}\right) + \beta_x\left(\gamma_{t+s} - \gamma_t\right) + \left(\delta_{t+s} - \delta_t\right) \\ &= \ln\left(\mu_{x,t}\right) + \beta_x\left(sc + \sum_{i=1}^{s} \varepsilon_{t+i}\right) + \left(\delta_{t+s} - \delta_t\right) \end{aligned}$$

That is

$$\mu_{x,t+s} = \mu_{x,t} \exp\left(\beta_x\left(sc + \sum_{i=1}^{s} \varepsilon_{t+i}\right) + \left(\delta_{t+s} - \delta_t\right)\right) \tag{4.38}$$

Since about 95 percent of the variance in the long-term forecasts is generated by the innovation of the latent factor γ_t, as reported by Lee and Carter (1992), one can simplify the forecast formula of $\mu_{x,t+s}$ as

$$\mu_{x,t+s} = \mu_{x,t} \exp\left(\beta_x \left(sc + \sum_{i=1}^{s} \varepsilon_{t+i}\right)\right) \tag{4.39}$$

The survival probability of the x-year-old over one year, assuming that the force of mortality is constant during the year $\mu_{x+u,t+u} = \mu_{x,t}$ $(0 \leq u < 1)$, is given by

$$p_{x,t} = p_{[x_0]+t,t} = \exp\left(-\mu_{x,t}\right) \tag{4.40}$$

The survival probability of the x-year-old over τ years is given by

$$_{\tau}p_{x,t} = \exp\left(-\sum_{i=1}^{\tau} \mu_{x+i,t+i}\right) \tag{4.41}$$

4.9.1 Estimation Procedure of LC92 Model

Let Y denote the matrix of log mortality rates, with each row for each age group $\ln \mu_x$ for N historical observations. We first construct a demeaned matrix of log mortalities, $X = Y - \alpha_x \iota$, where α_x is the mean value of $\ln \mu_x$, and ι is a row vector of ones. Then, as proposed by Lee and Carter (1992), we can use Singular Value Decomposition (SVD) to estimate the latent factor γ_t and the age-specific sensitivity β_x. $X = USV'$. Since the first singular value is significantly larger than other singular values, one can use one factor to approximate the log of force of mortality, as proposed by Lee and Carter (1992). β_x is the first column of U (multiplied by -1 to keep γ_t a downward sloping trend), and γ_t is the first element of S times the first column of V (multiplied by -1 to keep γ_t a downward sloping trend). The straightforward estimations of the drift parameter c, the variance of the innovation of the latent factor, and the variance of the estimated c are given by:

$$\begin{aligned}
\widehat{c} &= \frac{1}{N-1}\sum_{n=2}^{N}\triangle\gamma_n = \frac{1}{N-1}\left(\gamma_N - \gamma_1\right) \\
\widehat{\sigma}^2_{\varepsilon} &= \frac{1}{N-1}\sum_{n=2}^{N}\widehat{\varepsilon}^2_n = \frac{1}{N-1}\sum_{n=2}^{N}\left(\triangle\gamma_n - \widehat{c}\right)^2 \\
\sigma(\widehat{c}) &= \frac{\widehat{\sigma}_{\varepsilon}}{\sqrt{N-1}}
\end{aligned}$$

The estimated latent process (in the United Kingdom) is the following, including two temporary shocks captured by a 'WWI' dummy and a 'WWII' dummy:

$$\gamma_t^{UK} = -0.0725 + \gamma_{t-1}^{UK} + 0.65 * WWI_t + 1.9 * WWII_t + \varepsilon_t \tag{4.42}$$

with $\widehat{\sigma^{UK}}_{\varepsilon} = 0.169$.

The estimated latent process (in the Netherlands) is the following, including two temporary shocks captured by a 'flu' dummy and a 'WWII' dummy[15]:

$$\gamma_t^{NL} = -0.0748 + \gamma_{t-1}^{NL} + 1.85 * flu_t + 0.63 * WWII_t + \varepsilon_t \tag{4.43}$$

with $\widehat{\sigma^{NL}}_{\varepsilon} = 0.2176$. As pointed out in Lee-Carter (1992), the dummy variables only reduced the standard errors of the mortality forecast, but not the trend itself.

4.9.2 Simulation

The simulation steps:

1. simulate the latent factor for T periods according to $\gamma_{t+i} = \widehat{c} + \gamma_{t+i-1} + \varepsilon_i$, for $i = 1, ..., T$, where $\varepsilon_i \sim N\left(0, \widehat{\sigma}^2_{\varepsilon}\right)$, and $\gamma_t = \gamma_{2003}$ which is the last γ obtained from the estimation.

2. compute the force of mortality according to $\mu_{x+i,t+i} = \exp\left(\widehat{\alpha}_{x+i} + \widehat{\beta}_{x+i}\gamma_{t+i}\right)$, for $i = 1, ..., T$,

3. compute the survival probability of the x-year-old cohort according to

[15] The 'flu' dummy takes non-zero values for years { 1918 = 1; 1919 = -1} and zero elsewhere. The 'WWII' dummy takes non-zero values for years { 1940 = 1; 1941 = 1; 1942 = 1; 1943 = 1; 1944 = 1; 1945 = 1; 1946 = -6} and zero elsewhere.

${}_{\tau}p_{x,t} = \exp\left(-\sum_{i=1}^{\tau}\mu_{x+i,t+i}\right)$, for $\tau = 1, ..., T$

4. compute the survival index $S_t = N_t p_x$, where N is the initial size of the cohort.

5. repeat 1-4 steps for M times. As a by-product, calculate the mean, variance, and confidence interval of the forecasted survival probabilities and the survival index.

4.10 Appendix B: Derivation of the Results in Section 4

First introduce some notations. For any given value of b, we have the marginal utility as $u'(x) = \exp(-\alpha x)$, and the inverse function of marginal utility $u'(\cdot)$ as $I_v = -\frac{1}{\alpha}\ln(z)$. The inverse function of the utility function is denoted as $I_u = -\frac{1}{\alpha}\ln(-\alpha z)$.

Problem 1 (continued)

Set up the Lagrange

$$L = E\left[\int_0^T e^{-\delta t}u(D_t)\,dt + e^{-\delta T}u(W_T)\right] + \phi\left(W_0 - E\left[\int_0^T M_t D_t dt + M_T W_T\right]\right)$$

$$\begin{aligned}\frac{\partial L}{\partial D_t} &= 0 \quad\Rightarrow\quad e^{-\delta t}u'(D_t) = \phi M_t \\ \frac{\partial L}{\partial W_T} &= 0 \quad\Rightarrow\quad e^{-\delta T}u'(W_T) = \phi M_T\end{aligned}$$

$$\begin{aligned}D_t^* &= I_v\left(e^{\delta t}\phi M_t\right) = -\frac{1}{\alpha}\ln\left(e^{\delta t}\phi M_t\right) \\ &= -\frac{1}{\alpha}\delta t - \frac{1}{\alpha}\ln\phi - \frac{1}{\alpha}\ln M_t \\ W_T^* &= I_v\left(e^{\delta T}\phi M_T\right) = -\frac{1}{\alpha}\ln\left(e^{\delta T}\phi M_T\right)\end{aligned}$$

Plug into the budget constraint and the indirect utility function, we have

$$\begin{aligned} W_0 &= E\left[\int_0^T M_t D_t^* dt + M_T W_T^*\right] && (4.44)\\ &= -\frac{1}{\alpha}E\left[\int_0^T M_t \ln\left(e^{\delta t}\phi M_t\right)dt + M_T \ln\left(e^{\delta T}\phi M_T\right)\right] && (4.45) \end{aligned}$$

$$\begin{aligned} V_0 &= E\left[\int_0^T e^{-\delta t}\left(-\frac{1}{\alpha}\exp\left(-\alpha D_t^*\right)\right)dt + e^{-\delta T}u\left(W_T^*\right)\right] && (4.46)\\ &= -\frac{1}{\alpha}\phi E\left[\int_0^T M_t dt + M_T\right] && (4.47) \end{aligned}$$

Problem 2 (continued)

Set up the Lagrange

$$\begin{aligned} L &= E\left[\int_0^T e^{-\delta t}u\left(D_t^\pi + E[S_t] - S_t\right)dt + e^{-\delta T}u\left(W_T^\pi\right)\right]\\ &\quad +\phi^\pi\left(W_0 + \pi - E\left[\int_0^T M_t D_t^\pi dt + M_T W_T^\pi\right]\right) \end{aligned}$$

Since longevity risk, $E[S_t] - S_t$, cannot be hedged in the modelled financial market, the optimal strategy, D_t^π, is independent from $E[S_t] - S_t$. Under the assumed preference (4.10), the above Lagrange can be rewritten as

$$\begin{aligned} L &= -\frac{1}{\alpha}E\left[\int_0^T e^{-\delta t}\exp\left(-\alpha D_t^\pi\right)E\left[\exp\left(-\alpha\left(E[S_t] - S_t\right)\right)\right]dt + e^{-\delta T}u\left(W_T^\pi\right)\right]\\ &\quad +\phi^\pi\left(W_0 + \pi - E\left[\int_0^T M_t D_t^\pi dt + M_T W_T^\pi\right]\right)\\ &= E\left[\int_0^T e^{-\delta t}u\left(D_t^\pi\right)G_t dt + e^{-\delta T}u\left(W_T^\pi\right)\right]\\ &\quad +\phi^\pi\left(W_0 + \pi - E\left[\int_0^T M_t D_t^\pi dt + M_T W_T^\pi\right]\right) \end{aligned}$$

where G_t denotes

$$G_t \equiv E\left[\exp\left(-\alpha\left(E[S_t]-S_t\right)\right)\right]$$

and α is the shorthand notation for $\alpha(W_0)=\overline{\alpha}W_0^{-b}$. G_t can be seen as a function of the certainty equivalent of $E[S_t]-S_t$.

The optimal dividend strategy can be found by

$$\begin{aligned}
\frac{\partial L}{\partial D_t^{\pi}} &= 0 \quad \Rightarrow \quad e^{-\delta t}u'\left(D_t^{*\pi}\right)G_t=\phi^{\pi}M_t \\
\frac{\partial L}{\partial W_T^{\pi}} &= 0 \quad \Rightarrow \quad e^{-\delta T}u'\left(W_T^{\pi}\right)=\phi^{\pi}M_T
\end{aligned}$$

$$D_t^{*\pi} = I_v\left(\frac{e^{\delta t}\phi^{\pi}M_t}{G_t}\right)=-\frac{1}{\alpha}\left[\ln\left(e^{\delta t}\phi^{\pi}M_t\right)-\ln G_t\right] \tag{4.48}$$

$$W_T^{*\pi} = I_v\left(e^{\delta T}\phi^{\pi}M_T\right)=-\frac{1}{\alpha}\ln\left(e^{\delta T}\phi^{\pi}M_T\right) \tag{4.49}$$

Plug into the budget constraint and the indirect utility function

$$\begin{aligned}
W_0+\pi &= E\left[\int_0^T M_t D_t^{*\pi}dt+M_T W_T^{*\pi}\right] \\
&= -\frac{1}{\alpha}E\left[\int_0^T M_t\ln\left(e^{\delta t}\phi^{\pi}M_t\right)dt+M_T\ln\left(e^{\delta T}\phi^{\pi}M_T\right)\right] && (4.50)\\
&\quad +\frac{1}{\alpha}E\left[\int_0^T M_t\ln G_t dt\right] && (4.51)
\end{aligned}$$

$$\begin{aligned}
V_0^{*\pi} &= E\left[\int_0^T e^{-\delta t}\left(-\frac{1}{\alpha}\exp\left(-\alpha D_t^{*\pi}\right)\right)G_t dt+e^{-\delta T}u\left(W_T^{*\pi}\right)\right] && (4.52)\\
&= -\frac{1}{\alpha}\phi^{\pi}E\left[\int_0^T M_t dt+M_T\right] && (4.53)
\end{aligned}$$

Equalizing the two indirect utilities (4.46) and (4.52), $V_0=V_0^{*\pi}$, we find $\phi=\phi^{\pi}$.

Comparing the two budget constraints (4.44) and (4.50), with $\phi = \phi^{\pi}$, we have

$$\begin{aligned} W_0 &= -\frac{1}{\alpha} E\left[\int_0^T M_t \ln\left(e^{\delta t}\phi^{\pi} M_t\right) dt + M_T \ln\left(e^{\delta T}\phi^{\pi} M_T\right)\right] && (4.54)\\ W_0 + \pi &= -\frac{1}{\alpha} E\left[\int_0^T M_t \ln\left(e^{\delta t}\phi^{\pi} M_t\right) dt + M_T \ln\left(e^{\delta T}\phi^{\pi} M_T\right)\right] && (4.55)\\ &\quad +\frac{1}{\alpha} E\left[\int_0^T M_t \ln G_t dt\right] && (4.56) \end{aligned}$$

The difference between the two budget constraints gives the expression for longevity risk loading of the survival bond

$$\begin{aligned} \pi &= \frac{1}{\alpha} E\left[\int_0^T M_t \ln G_t dt\right] = \frac{1}{\alpha}\int_0^T E[M_t] \ln G_t dt && (4.57)\\ &= \frac{1}{\alpha}\int_0^T e^{-rt} \ln G_t dt. && (4.58) \end{aligned}$$

The risk loading is a present value of the certainty equivalent compensations $\frac{1}{\alpha}\ln G_t$ for the risks $S_t - E[S_t]$. This compensation is paid out as part of dividend in each period, as shown in the optimal dividends policy (4.48).

References

Allen, F. and Gale, D. (1997), Financial markets, intermediaries, and intertemporal smoothing, Journal of Political Economy, 105, 523-545.

Amromin, G., J. Huang and C. Sialm (2007), The Trade-off Between Mortgage Prepayments and Tax-Deferred Retirement Savings, Journal of Public Economics 91, p. 2014-2040

Antolin, P. and Blommestein, H. 2007, Governments and the Market for Longevity Indexed Bonds, OECD WP on insurance and private pensions, No. 4

Bader, L.N., and Gold, J. (2002), Reinventing pension actuarial science, Pension Section SOA, Volume 14, No. 2, January 2003, Pension Forum Book.

Bauer, D., and J. Russ (2006), Pricing Longevity Bonds Using Implied Survival Probabilities, working paper.

Benartzi, S., E. Peleg, and R.H. Thaler (2007), Choice Architecture and Retirement Saving Plans, working paper, Available at SSRN: http://ssrn.com/abstract=999420

Benzoni, L., P. Collin-Dufresne, and R.S. Goldstein (2007), Portfolio Choice Over the Life-Cycle When the Stock and Labor markets are Cointegrated, Journal of Finance, 2005 October, VOL. LXII, NO. 5, pp 2123-2167

Beshears, J., J.J. Choi, D. Laibson, and B.C. Madrian (2007), The Importance of Default Options for Retirement Saving Outcomes: Evidence from the United States, Harvard University, working paper

Beshears, J., J.J. Choi, D. Laibson, B.C. Madrian, and B. Weller (2008), Public Policy and Saving for Retirement: The "Autosave" Features of the Pension Protection Act of 2006, Harvard University, working paper

Blake D, Dowd, K, Cairns, A, Dawson, P (2006), Survivor Swaps, Journal of

Risk and Insurance 73 , 1-17

Blake D., A. Cairns, Dowd, K (2006a), Living with Mortality: Longevity Bonds and Other Mortality-linked Securities, British Actuarial Journal, Volume 12, Number 1, 2006 , pp. 153-197

Blake D., Cairns, A, Dowd, K (2006b), Mortality-Dependent Financial Risk Measures, Insurance: Mathematics and Economics 38 , 427-440

Blake, D. (1998), Pension schemes as options on pension fund assets: implications for pension fund management, Insurance, Mathematics and Economics, vol. 23, pp 263-286.

Bodie, Z., D. Mcleavey, L.B. Siegel (2007), The future of life-cycle saving and investing, The Research Foundation of CFA Institute, ISBN 978-0-943205-96-0

Bohn H. (2003): Intergenerational risk sharing and fiscal policy, working paper, University of California at Santa Barbara.

Cairns, A.J.G., Blake, D., and Dowd, K., (2005b), A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty, to appear in Journal of Risk and Insurance.

Cairns, A.J.G., Blake, D., and Dowd, K., (2006), Pricing Death: Frameworks for the Valuation and Securitization of Mortality Risk, to appear in ASTIN Bulletin, volume 36.1

Cairns, A.J.G., Blake, D., Dawson, P., and Dowd, K., (2005a), Pricing the Risk on Longevity Bonds. Life and Pensions October: 41-44

Calvet, L.E., J.Y. Campbell, and P. Sodini (2007a), Down or Out: Assessing the Welfare Costs of Household Investment Mistakes, The Journal of Political Economy, 2007, vol. 115, no. 5, 707-747.

Calvet, L.E., J.Y. Campbell, and P. Sodini (2007b), portfolio Rebalancing by Individual Investors

Campbell, J.Y., and L.M. Viceira (2002), Strategic Asset Allocation, Oxford University Press.

Carroll, C.D. (1992), The Buffer Stock Theory of Saving: Some Macroeconomic Evidence, Brookings Papers on Economic Activity, 2, 61-135.

Carroll, C.D. (1994), How Does Future Income Affect Current Consumption? Quarterly Journal of Economics, 109, 111-147.

Carroll, C.D. (1997), The Buffer Stock Saving and the Life Cycle / Permanent

Income Hypothesis, Quarterly Journal of Economics, 107, 1-56.

Carroll, C.D. (2006), The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems, Economic Letters, 312-320.

Carroll, C.D. (2007), Lecture notes on solution methods for microeconomic dynamic stochastic optimization problems, working paper, Department of Economics, Johns Hopkins University.

Chapman, R.J., Gordon, T.J., and Speed, C.A. (2001), Pensions, funding and risk, British Actuarial Journal, vol. 7, nr 4, pp. 605-663.

Chen, A. A. Pelsser, and M. Vellekoop (2007), Approximate solutions for indifference pricing with general utility functions, working paper, University of Amsterdam

Choi, J.J., D. Laibson, B. Madrian, A. Metrick (2004), For Better or For Worse: Default Effects and 401(k) Savings Behavior, in D.A. Wise, ed., Perspectives on the Economics of Aging, Chicago, University of Chicago Press, 81-121

Cocco, J.F., (2005), Portfolio Choice in the Presence of Housing, The Review of Finance Studies, 18(2), 535-567.

Cocco, J.F., F.J. Gomes and P.J. Maenhout (2005), Consumption and portfolio choice over the life cycle, The Review of Financial Studies, 18(2), 491-533.

Cox, S.H. and Y. Lin (2004), Natural hedging of life and annuity mortality risks. In Proceedings of the 14th International AFIR Colloquium, Boston, pp483-507.

Cui, J. and F.C.J.M. de Jong and E.H.M. Ponds (2008), Intergenerational risk sharing within funded pension schemes, Tilburg University, working paper

Cui, J. (2008a), Longevity risk pricing, Netspar Discussion paper. DP 2008-001

Cui, J. (2008b), DC pension plan defaults and individual welfare, Tilburg University, working paper

Dahl, M.H. and T. Moller (2006), Valuation and hedging of life insurance liabilities with systematic mortality risk, Insurance: Mathematics and Economics 39 (2006), 193-217

Dahl, M.H., (2004), Stochastic mortality in life insurance: market reserves and mortality-linked insurance contracts. Insurance: Mathematics and Economics 35, 113–136.

Dammon, R.M., C.S. Spatt and H.H. Zhang (2004), Optimal Asset Location and Allocation with Taxable and Tax-Deferred Investing, The Journal of Finance, Vol. LIX, No. 3, June 2004.

De Jong, F.C.J.M. (2007), Pension fund investments and the valuation of liabilities under conditional indexation, forthcoming Insurance: Mathematics and Economics.

De Jong, F.C.J.M. (2007), Valuation of pension liabilities in incomplete markets. working paper.

Diamond, P.A. and P.R. Orszag (2005), Saving social security, Journal of Economic perspectives 19, No. 2, Spring 2005, pp. 11-32

Enders W. and Lapan H. (1982), Social security taxation and intergenerational risk sharing, International Economic Review 23, pp 647-658.

Exley, C.J. (2004), Stakeholder interests alignment agency issues, paper presented at the Centre for Pension Management Colloquium October 5-6 2004, University of Toronto.

Fama, E.F., and French, K.R. (2002), The equity premium, Journal of Finance 57, 637-659.

Feldstein, M. (2005), Structural reform of social security, Journal of Economic Perspectives, vol 19, No. 2, Spring 2005, pp. 33-55

Fisher S. (1983), Welfare aspects of government issue of indexed bonds, in: R. Dornbusch and M. Simonsen (eds.): Inflation, debt and indexation, Cambridge, MIT Press.

Friedberg, L. and A. Webb (2005), Life is cheap: using mortality bonds to hedge aggregate mortality risk, CRR working paper 2005-13.

Gale D. (1990), The efficient design of public debt, in: R. Dornbusch and M. Draghi (eds): Public debt management: theory and history, Cambridge University Press.

Gollier, C. (2007), Intergenerational risk sharing and risk taking of a pension fund, working paper, University of Toulouse.

Gomes, F. and Michaelides, A. (2005), Optimal life-cycle asset allocation: understanding the empirical evidence, Journal of Finance, 60, pp. 869-904.

Gomes, F.J. and A. Michaelides (2005), Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence, The Journal of Finance, 60(2), 869-904.

Gomes, F.J., A. Michaelides and V. Polkovnichenko (2008), Optimal Savings with Taxable and Tax-Deferred Accounts, working paper

Gomes, F.J., L.J. Kotlikoff and L.M. Viceira (2008), Optimal Life-Cycle Investing with Flexible Labor Supply: a Welfare Analysis of Life-Cycle Funds, NBER working paper W13966

Gordon, R.H., and Varian, H.R. (1988), Intergenerational risk sharing, Journal of Public Economics, vol 14, pp. 1-29.

Gormley, T. H. Liu and G. Zhou (2007), limited participation and consumption-saving puzzles: A simple explanation, working paper, Washington University in St Louis.

Gourinchas, P., and J.A. Parker (2002), Consumption over the Life Cycle, Econometrica 71(1), 47-89.

Hari, N. (2006), Modeling Mortality: Empirical Studies on the Effect of Mortality on Annuity Markets, PhD thesis, Tilburg University

Hari, N., A. De Waegenaere, B. Melenberg, and T.E. Nijman (2008a), Estimating the Term Structure of Mortality, Insurance: Mathematics and Economics, forthcoming

Hari, N., A. De Waegenaere, B. Melenberg, and T.E. Nijman (2008b),Longevity Risk in Portfolios of Pension Annuities, Insurance: Mathematics and Economics, forthcoming

Henderson, V. (2002), Valuation of claims on nontraded assets using utility maximization, Mathematical Finance 12, 351-373

Kaas, R., M.J. Goovaerts, J. Dhaene and M. Denuit (2001). Modern Actuarial Risk Theory. Kluwer Academic Publishers, Dordrecht.

Krueger, D., and F. Kubler (2006), Intergenerational Risk Sharing via Social Security?, The American Economic Review, vol. 96, No. 3, June 2006, pp. 737-755.

Lee, R.D. and L.R. Carter (1992), Modeling and forecasting U.S. Mortality, Journal of the American Statistical Association, Vol 87, No. 419 (sep 1992), 659-671.

Lin, Y. and S.H. Cox (2005), Securitization of mortality risks in life annuities. The journal of risk and insurance, 72, 227-252

Lusardi, A. and O.S. Mitchell (2006), Baby Boomer Retirement Security: the

Roles of Planning, Financial Literacy, and Housing Wealth, University of Michigan Working paper WP 2006-114

Lusardi, A. and O.S. Mitchell (2007), Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education Programs, The Pension Research Council, working paper, WP2007-04

Lynch, A.W. and S. Tan (2006), Labor income dynamics at business cycle frequencies: Implications for portfolio choice, Working paper, New York University.

Merton, R.C. (1969), Lifetime portfolio selection under uncertainty: The continuous-time case, Review of Economics and Statistics, 51, August, 1969, 247-57.

Merton, R.C. (1971)

Merton, R.C. (1983), On the role of social security as a means for efficient risk sharing in an economy where human capital is not tradeable, in Bodie Z. and Shoven J.B (eds), Financial aspects of the United States pension system, Chicago.

Milevsky, M.A., S.D. Promislow and V.R. Young (2005), Financial valuation of mortality risk via the instantaneous Sharpe ratio, working paper

Milevsky, M.A., S.D. Promislow and V.R. Young (2006), Killing the law of large numbers: mortality risk premia and the Sharpe ratio, working paper

Moore, Kristen S. and Virginia R. Young (2003), Pricing equity-linked pure endowments via the principle of equivalent utility, Insurance: Mathematics and Economics 33 (2003) 497-516

Mottola and Utkus (2007), Red, Yellow, and Green: A Taxonomy of 401(k) Portfolio Choices, The Pension Research Council, working paper, WP2007-14

Musiela, Marek and Thaleia Zariphopoulou (2004), An example of indifference prices under exponential preferences, Finance and Stochastics 8 (2004) 229-239

OECD (2006), Pension Markets In Focus, October 2006, Issue 3

Oeppen, Jim and James W.Vaupel (2002), Broken Limits to Life Expectancy, Science 296, no. 5570 (2002): 1029-31.

Pelsser, A. (2005), Market-Consistent Valuation of Insurance Liabilities. working paper.

Ponds, E.H.M. (2003), Pension funds and value-based generational accounting, Journal of Pension Economics and Finance, vol. 2, nr. 3, pp. 295-325.

Ponds, E.H.M., and B. van Riel (2007), The recent evolution of pension funds in the Netherlands: the trend to hybrid DB-DC plans and beyond, working paper

2007-9, Center for Retirment Research at Boston College.

Poterba, J., Rauh, J., Venti, S. and Wise, D. (2005), Lifecycle asset allocation strategies and the distribution of 401(k) retirement wealth, working paper MIT

Scholz, J.K., A. Seshadri, and S. Khitatrakun (2006), Are Americans Saving "Optimally" for Retirement?, Journal of Political Economy, 2006, vol. 114, no. 4. pp 607-643

Schrager, D.F. (2006), Affine stochastic mortality, Insurance: Mathematics and Economics 38 (2006), 81-97

Sharpe, W.F. and L.G. Tint (1990), Liabilities: A New Approach, Journal of Portfolio Management,

Shell, Karl (1971), Notes on the economics of infinity, Journal of Political Economy, 79, 5, 1002-1011.

Shiller, R.J. (1999), Social security and institutions for intergenerational, intragenerational and international risk sharing, Carnegie-Rochester Conference Series on Public Policy, vol. 50, pp 165-204.

Svensson, Lars E.O., and Ingrid Werner (1993), Nontraded assets in incomplete markets, European Economic Review 37, 1149-1161.

Teulings, C., and C. de Vries (2006), Generational Accounting, Solidarity and Pension Losses, De Economist, vol. 154, no. 1, (March 2006), pp 63-83

Thaler, R. H. and S. Benartzi (2004), Save More Tomorrow^{TM}: Using Behavioral Economics to Increase Employee Saving, Journal of Political Economy, 2004, vol. 112, no. 1, pt.2, 164-187

Van Bommel, J. (2006), Intergenerational risk sharing and bank raids', working paper, University of Oxford.

Van Hemert, O. (2005), Optimal intergenerational risk sharing, UBS PRP Discussion Paper No. 37.

Viceira, L.M. (2001), Optimal Portfolio Choice for Long-horizon Investors with Nontradable Labor Income, Journal of Finance, Vol 56, Issues 2, pp 433-470

Viceira, L.M. (2007), Life-Cycle Funds, manuscript, Harvard University, forthcoming in Annamaria Lusardi, ed., Overcoming the saving slump: How to increase the effectiveness of financial education and saving programs, University of Chicago Press.

Wachter, J. and M. Jogo (2007), Why do household portfolio shares rise in wealth?, working paper, University of Pennsylvania, Electronic copy available at: http://ssrn.com/abstract=970953

Weiss L. (1979), The effects of money supply on economic welfare in the steady state, Econometrica 48, pp 565-576.

Young, V.R. (2004), Premium Principles, Encyclopedia of Actuarial Science, John Wiley & Sons. Ltd, 2004

Young, V.R. and Zariphopoulou, T. (2002) "Pricing Dynamic Insurance Risks Using the Principle of Equivalent Utility," Scandinavian Actuarial Journal, Vol. 2002(4), 246-279.

Nederlandse Samenvatting (Dutch Summary)

Door de stijging van de levensverwachting en de daling van de vruchtbaarheid in het verleden en in de nabije toekomst staan publieke pensioenstelsels (gefinancierd middels het omslagstelsel) onder toenemende druk, met als gevolg een verschuiving in de richting van kapitaalgedekte pensioensystemen als de belangrijkste bron van pensioeninkomsten. Tot het begin van deze eeuw waren traditionele kapitaalgedekte pensioensystemen voornamelijk collectieve defined-benefit (DB) regelingen, gegarandeerd door de overheid en de werkgever. Echter, de vergrijzing en nieuwe boekhoudregels op basis van marktwaardering van verplichtingen hebben als gevolg gehad dat de financieringskosten van deze traditionele DB regelingen dramatisch is geëscaleerd – en het is een worsteling geworden voor publieke en private sponsoren om aan hun verplichtingen te voldoen. Voornamelijk gedurende de crisis van 2001-2003 zijn veel DB regelingen ofwel niet meer toegankelijk voor nieuwe deelnemers, ofwel compleet bevroren. Overal ter wereld voelden kapitaalgedekte regelingen wereldwijd zich genoodzaakt om hun regelingen te hervormen.

In de afgelopen jaren zijn kapitaalgedekte pensioensystemen twee richtingen opgegaan. De eerste richting is de afschaffing van het collectieve karakter in de richting naar individuele pensioenregelingen waarbij de financiële risico´s volledig verschuiven van de sponsor naar de individuele deelnemer. De individuele defined-contribution (DC) regelingen zijn hier typische voorbeelden van. Echter, DC regelingen verschuiven niet alleen de risico´s naar de individuele deelnemer, maar confronteren de deelnemer ook met complexe investeringsbeslissingen. De tweede richting behoudt het collectieve karakter van kapitaal gedekte DB regelingen, maar verdeelt de financiële risico´s over alle stakeholders (gepensioneerden, werknemers en de sponsor). Verscheidene hybride collectieve regelingen zijn hiervan typische voorbeelden geworden. Hiermee komt de vraag naar boven hoe de risico´s van het pensioenfonds verdeeld dienen te worden onder de stakeholders. Is de ene richting beter dan de andere richting, of kunnen beide richtingen verbeterd worden? Dan zijn de vragen die dit proefschrift probeert te beantwoorden.

Dit proefschrift draagt bij aan ons economische begrip van de relatieve kracht en zwakheden van zowel collectieve als individuele pensioensystemen en introduceert verschillende voorstellen voor de verbetering van kapitaalgedekte pensioensystemen. In een notendop ligt de economische toegevoegde waarde van collectieve systemen in efficiënte risicodeling, en het ontbreken daarvan is tevens de zwakte van individuele systemen. De kracht van individuele systemen ligt in de mogelijkheden af te stemmen op de heterogene preferenties van deelnemers. De volgende generatie pensioensystemen zou daarom de krachten van beide werelden moeten combineren. Dat wil zeggen dat de intergenerationele risicodeling uit collectieve regelingen alsmede de levensloopkarakteristieken uit individuele regelingen behouden blijven. In de resterende hoofdstukken zullen deze argumenten verder uitgewerkt worden.

Hoofdstuk 2 ("Intergenerational risk-sharing in funded pension schemes") van dit proefschrift legt de focus op een uniek aspect van collectieve pensioenregelingen: de mogelijkheid tot intergenerationele risicodeling (IRD) tussen werknemers en gepensioneerden in dezelfde pensioenregeling. De lange termijn karakteristieken van collectieve pensioenregelingen maken het mogelijk dat risico's kunnen worden gedeeld tussen vele generaties. Collectieve pensioenregelingen gebaseerd op IRD hebben daarmee bredere mogelijkheden tot het verdelen van en verzekeren tegen risico's in vergelijking tot individuele systemen die zich beperken tot de onzekerheden van een enkel individu. We laten zien dat goed georganiseerde intergenerationele risicodeling binnen realistisch ontworpen collectieve systemen welvaartsverhogend kunnen zijn in vergelijking tot optimaal individueel pensioensparen. Deze bevinding heeft belangrijke implicaties voor pensioenhervormingen in zowel collectieve als individuele pensioenregelingen. Bij zulke hervormingen dient de mogelijkheid tot intergenerationele risicodeling behouden te blijven en versterkt te worden, en dus niet te worden afgeschaft. Onze casus laat zien dat IRD implementeerbaar is in realistische pensioenfondsen.

Hoofdstuk 2 is sterk gerelateerd aan drie onderzoeksgebieden: Asset Liability Management (ALM), Intergenerationele risicodeling (IRD) en Contingent Claim Analysis. Het traditionele ALM raamwerk neemt het DB beleid met betrekking tot contributies en uitbetalingen als gegeven en richt zich op de strategische beleggingsbeslissingen. Onze aanpak wijkt af van dit raamwerk, in de zin dat we het ontwerp van contributie- en uitbetalingsbeleid integreren in het ALM raamwerk; het pensioenbeleid (i.e. de regels tot het verdelen van risico's) worden daarmee gezamenlijk met het beleggingsbeleid geoptimaliseerd. We laten zien dat het pensioenbeleid niet alleen invloed heft op het beleggingsbeleid, maar ook op het welvaartsniveau van de deelnemers. Voortbouwend op de IRD literatuur, die zich voornamelijk richt op publieke financiering, analyseren we of IRD wenselijk is in kapitaalgedekte pensioenregelingen – en indien wenselijk hoe risico's optimaal verdeeld dienen te worden. Onze waarderingsmethode is tevens consistent met de waardering van verplichten op marktwaarde. Contingente claim waarderingstechnieken maken het mogelijk om de marktwaarde van IRS te ontleden in call en put opties in het bezit van de verschillende generaties. We laten zien dat, ex ante, de marktwaarde van de call optie gelijk is aan de marktwaarde van de put optie en dat IRS kan leiden tot een verhoging van de welvaart van deelnemers van collectieven pensioenregelingen.

Echter, het moge duidelijk zijn dat collectieve pensioenregelingen met IRS ook niet perfect zijn. Een belangrijk punt van kritiek is dat het uniforme beleid van deze regelingen zichzelf niet leent voor het heterogene profiel van de deelnemers. Een dergelijke afstemming op het profiel van de individuele deelnemer is exact de kracht van een (ideale) individuele pensioenregeling. Hoofdstuk 3 bestudeerd derhalve het ontwerp van een individuele DC pensioenvoorziening in een realistisch gekalibreerd levensloop raamwerk.

De studie naar individuele DC pensioenregelingen in Hoofdstuk 3 ("DC Pension Plan Defaults and Individual Welfare") dient een tweetal doelen. Ten eerste zijn de optimale contributie en investeringsregels belangrijk voor aanpassingen in het ontwerp van collectieve pensioenregelingen (bijvoorbeeld de opname van een levensloopprofiel in collectieve beleidsregels). Dit leidt tot een nieuw onderzoektraject op het gebied van inkomens- en leeftijdsafhankelijke beleidsregels voor collectieve contracten. Ten tweede bevatten de resulterende optimale levensloopstrategieën nuttige indicatoren voor het ontwerp van de default parameters van individuele regelingen. Theoretisch gezien zijn deze default parameters irrelevant, aangezien elke deelnemers de vrijheid heeft om de beleidsparameters naar eigen inzicht te kiezen. In de praktijk echter observeren we dat de meeste mensen eenvoudigweg de default paramters volgen, die doorgaans leeftijdsonafhankelijk zijn. Gegeven dat het default beleid van grote invloed is op het welvaartsniveau van deelnemers, introduceert Hoofdstuk 3 een leeftijdsafhankelijke default beleid met betrekking tot contributies en investeringen. We laten zien dat eenvoudige leeftijdsafhankelijke defaults een grote welvaartsverbetering voor deelnemers als gevolg heeft (ten opzichte van het huidige leeftijdsonafhankelijke default beleid).

Hoofdstuk 3 bouwt voort op de levensloopliteratuur waarin consumptie en beleggingsbeslissingen over de levensloop geoptimaliseerd worden. Daarnaast is het hoofdstuk gerelateerd aan de paper van Gomes, Michaelides and Polkovnichenko (2006), waarin optimaal contributie- en beleggingsbeleid wordt geanalyseerd voor rationele individuen voor het geval met belastbare DC besparingen alsmede het geval waarbij de belasting op DC besparingen uitgesteld is. De modellering in Hoofstuk 3 heeft een vergelijkbare opzet, maar met focus op het optimale leeftijdafhankelijke default contributie- en investeringsbeleid. Ook Gomes, Kotlikoff and Viceira (2008) maken een welvaartsvergelijking tussen eenvoudige defaults, maar zij bestuderen slechts de defaults voor beleggingsbeleid. We vinden dat de default met betrekking tot de contributies (of besparingen) een grotere impact heeft op het welvaartsniveau van het individu dan de keuzes met betrekking tot het beleggingsbeleid.

Een vereenvoudiging die in hoofdstuk 3 gemaakt wordt heeft betrekking op de individuele strategie tot het kopen van annuïteiten. Hoewel de keuze tot het kopen van annuïteiten vereenvoudigd is op het individuele niveau, gaat Hoofdstuk 4 in op het risicomanagement van sterfterisico op het geaggregeerde niveau. Trends in sterftecijfers en de hieraan gerelateerde onzekerheden sturen in belangrijke mate de globale vergrijzing met als gevolg de massale hervormingen in de pensioenwereld. Zowel collectieve pensioenfondsen alsmede de verstrekkers van annuïteiten aan individuele deelnemers aan DC regelingen zijn blootgesteld aan sterfterisico. In een aantal bekende gevallen heeft dit geleid tot aanzienlijke solvabiliteitsproblemen voor pensioenfondsen en annuïteitverstrekkers.

Een effectieve manier op sterfterisico te beheersen is het risico door te geven aan de financiële markt in brede zin via longevity linked securities. Echter, als gevolg van de moeilijkheden met betrekking tot het prijzen van deze producten worden ze niet verhandeld op financiële markten.

Hoofdstuk 4 ("Longevity Risk Pricing") introduceert een nieuwe methode voor het prijzen van risicopremies op sterfterisico, met als doel om de obstakels van het incomplete markt raamwerk te overkomen. We laten zien dat, als gevolg van de incompleetheid van de markt, de hoogte van de risicopremie afhankelijk is van het uitbetaligsschema van het product. Verder laten we zien dat de financiële kracht van de koper en verkoper van de sterfteverzekering, de beschikbaarheid van natuurlijke beschermingsmogelijkheden en de aanwezigheid van basisrisico een significante invloed kunnen hebben op de hoogte van de risicopremie.

Hoofdstuk 4 combineert de literatuur op het gebied van het prijzen van financiële producten in een incomplete markt met de literatuur op het gebied van de stochastische modellering van sterfterisico om de moeilijkheden met het prijzen van innovatieve producten gebaseerd op sterfterisico te overkomen.

Op basis van deze bevindingen geeft dit proefschrift een aantal indicaties voor richtingen voor de verbeteringen van pensioensystemen in de toekomst. De volgende generatie pensioensystemen zou de krachten van collectieve en individuele regelingen kunnen combineren. Hoe kan dit worden bewerkstelligd? De DC regelingen kunnen bijvoorbeeld componenten voor winst- en risicodeling kunnen toevoegen om IRD te repliceren. De collectieve regelingen zouden levensloop-karakteristieken kunnen opnemen in hun beleid. Concluderend, om de pensioenzekerheid in een vergrijzende samenleving te kunnen blijven waarborgen zullen aanpassingen in het ontwerp van pensioencontracten en innovaties in financiële markten integraal moeten verlopen.